ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

Dottorato di Ricerca in Geofisica
Ciclo XXIV

Settore Concorsuale di afferenza: 04/A4
Settore Scientifico Disciplinare di afferenza: GEO/10

Spectral-Element and Adjoint 3D Full-Wave Tomography for the Lithosphere of Central Italy

Ph.D. Thesis of:
Federica Magnoni

Tutor:
Dr. Alberto Michelini

Co-Tutor:
Dr. Emanuele Casarotti

Ph.D. Coordinator:
Prof. Michele Dragoni

Esame finale anno 2012

Abstract

The primary objective of this thesis is to obtain a better understanding of the 3D velocity structure of the lithosphere in central Italy. To this end, I adopted the Spectral-Element Method to perform accurate numerical simulations of the complex wavefields generated by the 2009 M_w 6.3 LAquila event and by its foreshocks and aftershocks together with some additional events within our target region. For the mainshock, the source was represented by a finite fault and different models for central Italy, both 1D and 3D, were tested. Surface topography, attenuation and Moho discontinuity were also accounted for. Three-component synthetic waveforms were compared to the corresponding recorded data. The results of these analyses show that 3D models, including all the known structural heterogeneities in the region, are essential to accurately reproduce waveform propagation. They allow to capture features of the seismograms, mainly related to topography or to low wavespeed areas, and, combined with a finite fault model, result into a favorable match between data and synthetics for frequencies up to ∼0.5 Hz. We also obtained peak ground velocity maps, that provide valuable information for seismic hazard assessment. The remaining differences between data and synthetics led us to take advantage of SEM combined with an adjoint method to iteratively improve the available 3D structure model for central Italy. A total of 63 events and 52 stations in the region were considered. We performed five iterations of the tomographic inversion, by calculating the misfit function gradient - necessary for the model update - from adjoint sensitivity kernels, constructed using only two simulations for each event. Our last updated model features a reduced traveltime misfit function and improved agreement between data and synthetics, although further iterations, as well as refined source solutions, are necessary to obtain a new reference 3D model for central Italy tomography.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview of the problem

The final aim of this work is to obtain an improved 3D model for the velocity structure of central Italy by performing an iterative tomographic inversion that exploits full observed waveforms. As a starting point, one needs to be able to reproduce the seismic wavefield generated in the region by the recorded events.

Accurate modeling of seismic wave propagation in realistic 3D heterogeneous media represents a paramount task in seismology. The wavefield produced by an earthquake carries a great amount of extremely valuable information. Building up a detailed, forward model is crucial in view of predicting large earthquake ground motion and assessing seismic hazard. Moreover, comparing accurately simulated waveforms to observed ones represents a preliminary fundamental step in order to address seismological inverse problems. For 1D Earth models, one can use semi-analytical techniques to calculate the wavefield generated by a point source (Aki & Richards (3), 1980; Kennett (96), 1983; Dahlen & Tromp (51), 1998). However, when complex 3D models (and, possibly, sources of finite size) are involved, suitable numerical techniques are required to solve the seismic wave equation of motion. The computational resources are tremendously increasing at present. Thus, developing performant numerical methods and softwares that exploit this power to calculate accurate 3D synthetic seismograms is a task that attracts the attention of several research groups worldwide. It is worth mentioning, e.g., the European project SPICE (Seismic Wave Propagation and Imaging in Complex Media: A European Network, 2004-2007), that aimed at producing computational

tools for seismic wave propagation, earthquake motion, and seismic imaging; or another European project, QUEST (QUantitative estimation of Earth's seismic source and STructure), that followed SPICE and focuses on developing and exploiting powerful 3D simulation techniques to address both forward and inverse modeling; or also Global ShakeMovie (Tromp et al. (199), 2010), a Princeton University's global seismicity portal, that provides in near real time 1D and 3D synthetic seismograms for recent earthquakes.

Amongst the numerical approaches developed for seismic wavefield simulations, the most widely known is probably the finite difference (FD) method (Madariaga (123), 1976; Virieux (202), 1986), in which derivatives are approximated by using differences of functions at adjacent grid points. FD has been applied for local and regional 3D simulations (e.g., Olsen & Archuleta (150), 1996; Graves (79), 1996; Ohminato & Chouet (148), 1997; and more recently Kristek et al. (111), 2009), but it suffers from limitations when dealing with 3D full complexity, such as the presence of free surface topography (Robertsson (169), 1996; Ohminato & Chouet (148), 1997) or anisotropy (Igel et al. (89), 1995). Thus, for global models, only simplified geometries can be accounted for (e.g., Igel & Weber (87, 88), 1995, 1996). However, recent improvements are being developed, such as staggered-grid schemes for viscoelastic media with material interfaces, or optimally accurate schemes (e.g., Moczo et al. (141), 2007).

Other methods include, for example, spectral and pseudo-spectral techniques, that have been used for both regional (e.g, Carcione (23), 1994) and global (e.g., Furumura et al. (77), 1998) problems. They are characterized by high spatial accuracy and very low numerical dispersion with only a small number of grid points per wavelength (Chaljub et al. (32), 2007). Nevertheless, since they use global basis functions, they are limited to media with smooth variations, because numerical noise arises at sharp discontinuities. Moreover, simulations of surface waves suffer from lower accuracy with respect to body waves, as in FD methods.

Boundary element (e.g., Kawase (94), 1988) or boundary integral (e.g., Sánchez-Sesma & Campillo (171), 1991) techniques are well suited to handle realistic surface and interface topography. However, they are restricted to models with a finite number of homogeneous layers, and in 3D the increasing numerical costs require approximations that cause artefacts in the solutions (e.g., Bouchon et al. (20),1996).

Classical finite-element methods (FEMs) are characterized by a very high geometrical flexibility. The grid boundaries can coincide with major interfaces in the model; this allows one to tackle difficulties related to the implementation of complex geometries with topography and discontinuities. The technique has been successfully applied to seismic wave simulations in 3D sedimentary basins (e.g., Bao et al. (12), 1998). However, FEMs use low-degree polynomials to discretize functions on elements, which causes low spatial accuracy and high numerical dispersion (Marfurt (129), 1984). The resulting large linear systems need to be solved by using approximate, iterative routines. This increases the computational costs and complicates the implementation of the algorithm, especially on parallel computers with distributed memory.

Recently, the arbitrary high-order derivative discontinuous Galerkin method (ADER-DGM; e.g., Käser & Dumbser (92), 2006; Dumbser & Käser (65), 2006; Käser et al. (93), 2007; de la Puente et al. (52, 53), 2007, 2008) is attracting the attention of the seismological community, since it combines the flexibility of tetrahedric meshes as FEM, but it is more accurate and computationally efficient (e.g., Chaljub et al. (33), 2010). In particular, this numerical technique allows one to solve the governing PDE with arbitrarily high-order approximation in space and then in time, whose accuracy is automatically coupled to the space accuracy determined by the degree of the used basis polynomials. Moreover, all the calculations are performed locally on each element and its direct neighbors, thus the method can be very well implemented on parallel computers.

The numerical technique used in this work is the spectral-element method (SEM), a well established approach for accurately solving both forward and inverse seismic problems with reduced computational costs in terms of memory and CPU time (e.g., Tromp et al. (198), 2008). SEM was originally developed in computational fluid dynamics (Patera (156), 1984; Maday & Patera (124), 1989) and its adoption in seismology is only recent. Initially, the method was applied to simulate seismic wave propagation in 2D (Cohen et al. (46), 1993; Priolo et al. (166), 1994). At present, it is widely and successfully used for 3D global and regional problems (e.g., Komatitsch (98), 1997; Faccioli et al. (70), 1997; Komatitsch & Vilotte (104), 1998; Komatitsch & Tromp (100), 1999; Paolucci et al. (152), 1999; Chaljub (29), 2000; Komatitsch & Tromp (101, 102), 2002a,b; Capdeville et al. (22), 2003; Chaljub & Valette (30), 2004; Fichtner et al. (74),

2009a), and local-scale simulations in complex regions, for example southern California, USA, (Komatitsch et al. (109), 2004; Tape et al. (188, 189), 2009, 2010), Taipei, Taiwan, (Lee et al. (113, 114, 115), 2008, 2009a,b), Caracas, Venezuela (Delavaud et al. (55), 2006), Grenoble, France (Chaljub et al. (31), 2005; Stupazzini et al. (184), 2009; Chaljub et al. (33), 2010), northern Apennines, Italy (Stich & Morelli (182), 2007; Stich et al. (183), 2009) and central Italy (Magnoni et al. (126), in prep.). SEM combines the accuracy and exponential convergence rate of spectral methods with the flexibility of FEMs. The conjunction between the chosen basis functions and integration rule leads to a great simplification of the algorithm, that lends itself very well to be implemented on parallel architecture, drastically reducing the cost of calculations. Moreover, the method allows the incorporation of all the 3D complexities that influence seismic wave propagation, such as the 3D lateral heterogeneity, topography, bathymetry, major discontinuities in the model like faults and also fluid-solid boundaries, anisotropy, anelasticity, ellipticity, self-gravitation, rotation and effect of the oceans (Komatitsch & Vilotte (104), 1998; Komatitsch & Tromp (100), 1999; Komatitsch et al. (106, 107), 2000a,b; Komatitsch & Tromp (101, 102), 2002a,b; Chaljub et al. (32), 2007). As a consequence, SEM turns out to be highly powerful and performant in simulating forward seismic wavefields and this represents the key motivation of our choice. A more detailed presentation of the method can be found in section 2.2. Figure 1.1 shows an example of the wavefield generated by the 2009 L'Aquila (Italy) event and simulated by using the spectral-element code, SPECFEM3D 2.0, considered in this work (sections 2.4 and chapter 3).

Once a powerful method, such as SEM, is available to accurately simulate seismic waveforms, the paramount challenge is exploiting this capability to tackle the 3D inverse problem. The general aim is to minimize a measure of the difference between observed and synthetic waveforms, represented by a misfit function (section 2.3.1), while improving models of the Earth's structure and earthquakes' sources (e.g., Tape et al. (188, 189), 2009, 2010; Kim et al. (97), 2011).

In seismology, most traditional tomographic methods use ray theory to interpret traveltime or phase measurements between data and synthetics. Inversion methods that take into account finite-frequency effects by using so-called finite-frequency sensitivity kernels, instead of techniques based upon ray theory, have been introduced, e.g., by Marquering et al. (131) (1999), Zhao et al. (210) (2000) and Dahlen et al. (49) (2000).

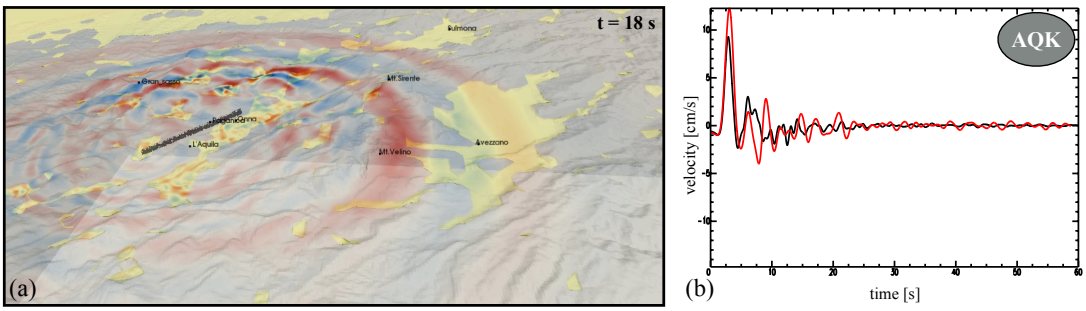


Figure 1.1: SEM simulations - (a) Z-component velocity wavefield (after 18 s) generated by the 2009 M_w 6.3 L'Aquila event, simulated by using SPECFEM3D 2.0 and a 3D model for the region. Red colors denote positive values and blue negative values. Yellow areas represent low velocity regions where the energy is trapped causing resonance effects. The black line is the fault trace. (b) Example of the comparison between Z-component velocity of our synthetic (red) and of the data (black) recorded at station AQK, 2 km far from the epicenter. The frequency range is 0.02-0.5 Hz. See section 3.7.

They considered 1D reference Earth models and calculated the sensitivity kernels using, respectively, surface-wave Green's functions, normal modes, and asymptotic, ray-based methods. This new kernel-based approach, to which seismic tomography is transitioning, recognizes that traveltime and amplitude anomalies of seismic waves are frequency-dependent, and incorporates traveltime effects associated with wavefront healing (e.g., Malcolm & Trampert (127), 2011). The kernels, in general, reflect the sensitivity of the waveforms to changes in the model parameters, hence leading the inversion. Other examples are Hung et al. (84) (2000) and Zhou et al. (212) (2004), which, for kernel calculations, used asymptotic, ray-based methods, Dahlen & Baig (50) (2002), that calculated kernels for body-wave amplitudes, Montelli et al. (142) (2004), which implemented kernels for compressional-wave tomography, or Zhao & Jordan (209) (2006), which used normal modes to compute global finite-frequency kernels for spherically symmetric models. Simple 3D traveltime kernels for phases like P and S are often called *banana-doughnut* kernels, due to their shape (e.g., Tromp et al. (197), 2005), and for 1D reference models their calculation is relatively inexpensive using techniques such as those cited above.

At present, the computational resources are dramatically growing, as well as the capabilities to simulate forward wavefields in complex media thanks to powerful methods

like SEM. Thus, performing tomographic inversions based upon 3D reference Earth models and fully 3D numerical simulations becomes feasible. Three-dimensional ray tracing techniques in 3D models have been used e.g. by Bijwaard & Spakman (14) (2000) to iteratively improve a global P-wave model. Zhao et al. (211) (2005), instead, used a FD method and obtained the 3D finite-frequency sensitivity kernels for 3D models by calculating and storing 3D Green's functions for every considered source and receiver as a function of space and time. This is a possible approach to construct kernels for 3D models and it would allow the calculation of both the first and second derivative of the misfit function to be used in the inversion (i.e., the gradient and the Hessian, respectively; section 2.3.1). However, for large 3D realistic applications, where numerous stations and earthquakes are involved, the storage requirements are almost unaffordable.

An alternative approach, followed in this work, to calculate 3D kernels for fully heterogeneous 3D media is based upon adjoint methods (e.g., Tromp et al. (197, 198), 2005, 2008; Liu & Tromp (118), 2006; Fichtner et al. (72, 73), 2006a,b. See section 2.3).

For any time window in a seismogram, e.g. a specific pick for a source-receiver pair, one could calculate the corresponding sensitivity kernel as the integrated interaction between the regular forward wavefield and an adjoint wavefield, whose source is located at the receiver and is based upon the difference between data and synthetics in the selected window (e.g., Tromp et al. (197), 2005). The gradient of traveltime or amplitude misfit functions may be expressed in terms of a weighted sum, over all source-receiver pairs, of these 3D banana-doughnut kernels, with weights determined by the corresponding traveltime or amplitude differences. Also the second derivatives of the misfit functions, i.e. the Hessian matrix, could be calculated from these phase-specific kernels (e.g., Tape et al. (187), 2007). Thus, in principle, to solve the inverse problem, one would need to perform one forward and one adjoint simulation for each measurement in a seismogram to construct the corresponding kernel, for a total of $(2N_{events}{*}N_{receivers}{*}N_{components}{*}N_{picks})$ simulations [1]. These calculations have a prohibitive computational cost for 3D realistic cases (e.g., Tromp et al. (198), 2008).

[1] N_{events} is the number of events, $N_{receivers}$ the number of receivers, $N_{components}$ the number of components for each receiver and N_{picks} is the number of picks or measurements contained in each component.

The main advantage of the adjoint approach is, thereby, that it is not necessary to calculate individual banana-doughnut kernels for each measurement (or time window), but one can directly calculate kernels that represent the overall sensitivity of the data, highlighting where the current 3D model is inadequate and needs to be modified (see section 2.3.3). For each event, these kernels are called *event kernels* and, in principle, they are weighted combinations of phase-specific banana-doughnut kernels, but, based upon the adjoint theory, they can be readily constructed by performing just two 3D simulations per event: one simulation for the forward wavefield and one for the adjoint wavefield, that now is generated by considering time-reversed signals at all the receivers as simultaneous sources (e.g., Tromp et al. (197), 2005; Tape et al. (187), 2007). Then, based upon Tarantola (190) (1984), the derivatives of the misfit function can be expressed as spatial integrals involving model perturbations and these 3D volumetric kernels, so an *adjoint calculation* can be thought as a means of explicitly determining the gradient of a misfit function needed for the inversion (Talagrand & Courtier (185), 1987). As a consequence of how the adjoint kernels are constructed, an essential advantage of the adjoint approach is that the number of required simulations to construct the gradient of a misfit function scales linearly only with the number of events, but it is independent of the number of stations and measurements. In addition, storing all the Green's functions is not necessary. It is worth noting, however, that in such an adjoint approach the calculation of the Hessian is practically unfeasible, since the single phase-specific kernels are not calculated due to computational costs. Thus, one needs to resort to iterative gradient-based procedures, e.g. conjugate gradient (e.g., Fletcher & Reeves (76), 1964) or steepest descent methods (e.g., Nolet (147), 1987), in order to minimize the misfit function and solve the inverse problem (e.g., Tarantola (190, 191, 192), 1984, 1987, 1988. See section 2.3.4).

Following Tape et al. (187) (2007), we refer to *adjoint tomography* as an inversion approach that uses only the misfit function gradient, calculated through adjoint methods, compared to a *classical tomography*, defined as a Newton inversion method based upon both the gradient and the Hessian of a misfit function (e.g., Woodhouse & Dziewonski (205), 1984; Ritsema et al. (168), 1999). The basics of the adjoint theory in seismology refer to Tarantola (190, 191, 192) (1984, 1987, 1988); early applications of the theory may be found in Gauthier et al. (78) (1986), Mora (144, 145) (1987, 1988), Crase et al. (48) (1990), Vasco et al. (201) (1995); Pratt et al. (165) (1998), Pratt (164)

(1999), Akçelik et al. (1, 2) (2002, 2003). The significant advantages of the method, as previously introduced, lead us to perform, in this work, a 3D adjoint tomography for central Italy lithosphere (chapter 4) by exploiting the powerful combination of SEM with the adjoint method, as shown, e.g., in Tromp et al. (198) (2008), Liu & Tromp (118) (2006), Tape et al. (187) (2007), Fichtner et al. (75) (2009b). A previous successful example we refer to is Tape et al. (188, 189) (2009, 2010), which presented a new 3D tomographic model for southern California crust after 16 iterations of the adjoint inversion procedure (combined with SEM simulations). In addition to tomographic inversions, adjoint methods can be also used to perform source inversions or joint source and structural inversions (e.g., Tromp et al. (197), 2005; Tape et al. (187), 2007; Kim et al. (97), 2011), but this is beyond the scope of this thesis.

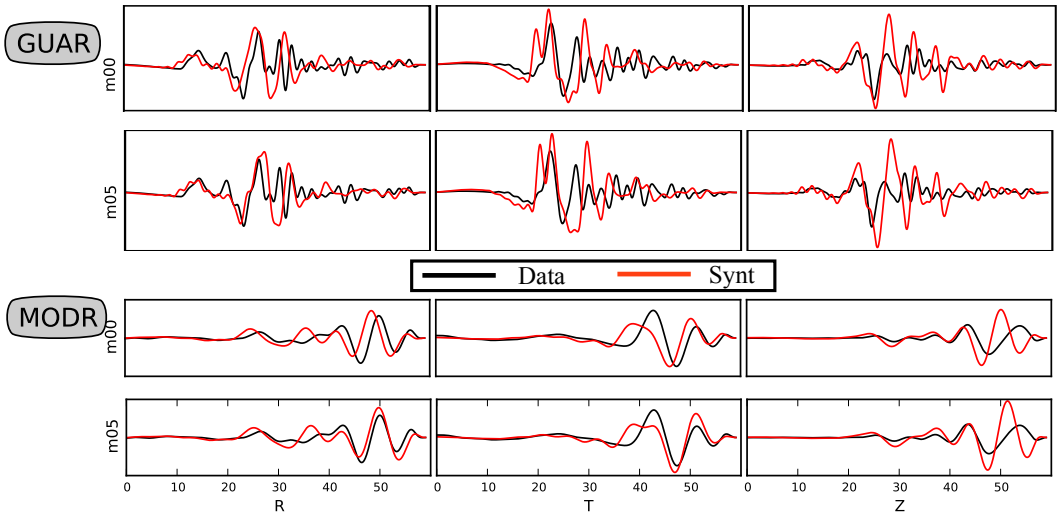


Figure 1.2: Adjoint tomography - Examples of the comparison between data (black) and synthetics (red) for the initial 3D velocity model of central Italy $\mathbf{m}^0$ and the updated model $\mathbf{m}^5$, obtained with 5 iterations of our adjoint tomographic inversion. The timeseries are for two different stations that recoded two different events. The top panel refers to the period range 2-20 s, the bottom panel to the range 6-20 s. See section 4.9.2

The following chapters and sections start with the introduction of the fundamental equations for the forward wavefield modeling (section 2.1). Then, section 2.2 will discuss in detail the spectral-element method used in this work. The basic theory of the adjoint approach for 3D tomographic inversions will be presented in section 2.3, after a brief outline of the inverse problem formulation, also in the classical approach.

In chapter 3 I will show the results obtained by using SEM for high frequency simulations of the wavefield generated by the 2009 M_w 6.3 L'Aquila earthquake, represented as a finite fault. The effects on ground motion simulations due to the introduction of models with topography and 3D heterogeneities will be discussed, comparing the synthetic seismograms to the observed data (see Fig. 1.1 (b) for a preview). This will also highlight the power of a numerical method like SEM, that is able to account for all these complexities, together with a finite source description, in modeling seismic wavefields. Moreover, this analysis provides valuable preliminaries for the study in chapter 4. In this section, the tomographic inversion for central Italy lithosphere, based on the adjoint method in combination with a steepest descent algorithm and SEM, will be presented. The last updated model, that we have obtained so far, results from 5 iterations of the inversion strategy. It represents an improved structure, compared to the initial model, that reduces the misfit between data and synthetics (see Fig. 1.2 for a preview) and features velocity patterns, that seem to be physically consistent and in agreement with recent studies about central Italy. This analysis yields promising results in view of obtaining a new 3D reference model of central Italy lithosphere, that requires however further improvements of the structure model with successive iterations, as well as an inversion also for seismic source parameters (e.g., Liu et al. (119), 2004, or Kim et al. (97), 2011; see Appendix A). Conclusions and future perspectives will be discussed in chapter 5.

Chapter 2

Forward and inverse modeling in seismology based upon Spectral-Element and Adjoint Methods

2.1 Basic theory of seismology

2.1.1 Constitutive relationships

Let us denote the displacement wavefield generated by an earthquake as $\mathbf{s}(\mathbf{x}, t)$, where $\mathbf{x}$=(x, y, z) is the position vector and t time.

For an elastic Earth the symmetric stress tensor $\mathbf{T}$ is linearly related to the displacement gradient $\nabla\mathbf{s}$ via the constitutive relationship

$$\mathbf{T} = \mathbf{c} : \nabla\mathbf{s}\,. \tag{2.1}$$

The fourth-order elastic tensor $\mathbf{c}$ or stiffness tensor denotes the elastic properties of the medium and in a general anisotropic case it has 21 independent components, due to its own symmetries.

In an isotropic, elastic Earth model the number of independent coefficients of $\mathbf{c}$ reduces to just two, the bulk modulus k and the shear modulus μ:

$$c_{jklm} = (\kappa - 2\mu/3)\,\delta_{jk}\delta_{lm} + \mu\,(\delta_{jl}\delta_{km} + \delta_{jm}\delta_{kl})\,, \tag{2.2}$$

and Hook's law 2.1 is expressed as

$$\mathbf{T} = (k - \frac{2}{3}\mu)\,\mathrm{tr}(\nabla\mathbf{s})\,\mathbf{I} + 2\mu\nabla\mathbf{s}\,. \tag{2.3}$$

In order to account for the anelasticity of the Earth, the constitutive equation for an attenuating medium relating the stress to the entire strain history (e.g., Komatitsch et al. (110), 2005) is

$$\mathbf{T}(t) = \int_{-\infty}^{t} \partial_t \mathbf{c}(t-t') : \nabla\mathbf{s}(t')\mathrm{d}t'\,. \tag{2.4}$$

In seismology, the quality factor Q is observed to be approximately constant over a wide range of frequencies. Such an absorption-band solid may be approximated using a series of L standard linear solids (Liu et al. (117), 1976). Usually, three of such linear solids are sufficient to accurately mimic an almost constant Q (Emmerich & Korn (69), 1987). In the Earth the attenuation is dominated by the shear quality factor $Q\mu$, that is several hundreds times smaller than the bulk quality factor Qk. Thus, in order to implement the attenuation, it is sufficient to model the time evolution of the isotropic shear modulus μ (Liu et al. (117), 1976)

$$\mu(t) = \mu_R \left[1 - \sum_{l=1}^{L} \left(1 - \frac{\tau_l^\epsilon}{\tau_l^\sigma}\right) e^{-t/\tau_l^\sigma}\right] H(t)\,, \tag{2.5}$$

where μ_R denotes the relaxed modulus, $H(t)$ the Heaviside function, and τ_l^σ and τ_l^ϵ the stress and strain relaxation times of the l-th standard linear solid.

Using equation 2.5 the constitutive relation 2.4 becomes

$$\mathbf{T} = \mathbf{c}_U : \nabla\mathbf{s} - \sum_{l=1}^{L} \mathbf{R}_l\,, \tag{2.6}$$

where the unrelaxed elastic tensor $\mathbf{c}_U$ is determined by the unrelaxed shear modulus

$$\mu_U = \mu_R \left[1 - \sum_{l=1}^{L} \left(1 - \frac{\tau_l^\epsilon}{\tau_l^\sigma}\right)\right]\,. \tag{2.7}$$

Thus, for each standard linear solid one has to solve the equation for the symmetric, traceless memory-variable tensor $\mathbf{R}_l$

$$\partial_t \mathbf{R}_l = -(\mathbf{R}_l - \delta\mu_l \mathbf{D})/\tau_l^\sigma\,, \tag{2.8}$$

where $\delta\mu_l$ is the modulus defect for the l-th standard linear solid

$$\delta\mu_l = -\mu_R(1 - \tau_l^\epsilon/\tau_l^\sigma)\,, \tag{2.9}$$

and $\mathbf{D}$ is the traceless strain deviator

$$\mathbf{D} = \frac{1}{2}[\nabla\mathbf{s} + (\nabla\mathbf{s})^T] - \frac{1}{3}(\nabla\cdot\mathbf{s})\,\mathbf{I}\,; \tag{2.10}$$

the superscript T denotes the transpose and $\mathbf{I}$ is the identity tensor.

2.1.2 Equation of motion

Assume a finite Earth model with volume Ω, free surface $\partial\Omega$, artificial absorbing boundary Γ and possible internal discontinuities. The unit outward normal to the boundary $\partial\Omega$+Γ is denoted by $\hat{\mathbf{n}}$, as well as the unit upward normal to any discontinuity.

In such an Earth model, the strong or differential formulation of the seismic wave equation that governs the displacement field $\mathbf{s}(\mathbf{x},t)$ associated to an earthquake is

$$\rho\,\partial_t^2\mathbf{s} = \nabla\cdot\mathbf{T} + \mathbf{f}\,, \tag{2.11}$$

where ρ is the distribution of mass density and the stress tensor $\mathbf{T}$ is related to the displacement gradient by one of the constitutive relations in section 2.1.1, depending on the considered Earth model.

The force $\mathbf{f}$ in equation 2.11 represents the earthquake source. Using a point source representation one can write $\mathbf{f}$ in terms of the moment tensor $\mathbf{M}$ as (Dahlen & Tromp (51), 1998)

$$\mathbf{f} = -\mathbf{M}\cdot\nabla\delta(\mathbf{x}-\mathbf{x}_s)\,S(t)\,, \tag{2.12}$$

where the point source location is $\mathbf{x}_s$, $\delta(\mathbf{x}-\mathbf{x}_s)$ is the Dirac delta distribution located at $\mathbf{x}_s$ and $S(t)$ the source-time function. A source of finite size, such as a finite fault plane Σ_s, may be modeled as a number of point sources, each with its own time history; thus, $\mathbf{f}$ can be expressed in terms of a moment-density tensor $\mathbf{m}$ as (e.g., Komatitsch et al. (110), 2005)

$$\mathbf{f} = -\mathbf{m}(\mathbf{x}_s,t)\cdot\nabla\delta(\mathbf{x}-\mathbf{x}_s) \qquad \text{on}\,\Sigma_s\,. \tag{2.13}$$

The equation of motion 2.11 is subject to the initial conditions

$$\mathbf{s}(\mathbf{x},0) = \mathbf{0}, \qquad \partial_t\mathbf{s}(\mathbf{x},0) = \mathbf{0}\,. \tag{2.14}$$

The boundary condition at the free surface requires that the traction $\hat{\mathbf{n}}\cdot\mathbf{T}$ vanishes on $\partial\Omega$, that is

$$\hat{\mathbf{n}}\cdot\mathbf{T} = \mathbf{0}\,. \tag{2.15}$$

On boundaries between solid materials, both traction $\hat{\mathbf{n}} \cdot \mathbf{T}$ and displacement $\mathbf{s}$ must be continuous; on fluid-solid boundaries traction $\hat{\mathbf{n}} \cdot \mathbf{T}$ and normal component of displacement $\hat{\mathbf{n}} \cdot \mathbf{s}$ must be continuous.

For regional and local simulations, the model is not the entire Earth and one needs to introduce fictitious boundaries Γ of the domain on which seismic energy must be absorbed to mimic a semi-infinite medium. An approximate absorbing boundary condition may be expressed as (Clayton & Engquist (44), 1977; Quarteroni et al. (167), 1998)

$$\hat{\mathbf{n}} \cdot \mathbf{T} = \rho \left[v_n (\hat{\mathbf{n}} \cdot \partial_t \mathbf{s}) \hat{\mathbf{n}} + v_1 (\hat{\mathbf{t}}_1 \cdot \partial_t \mathbf{s}) \hat{\mathbf{t}}_1 + v_2 (\hat{\mathbf{t}}_2 \cdot \partial_t \mathbf{s}) \hat{\mathbf{t}}_2 \right] . \tag{2.16}$$

$\hat{\mathbf{t}}_1$ and $\hat{\mathbf{t}}_2$ are orthogonal unit vectors tangential to Γ, v_n is the quasi-P wave speed of waves traveling in the $\hat{\mathbf{n}}$ direction, v_1 is the quasi-S wave speed of waves polarized in the $\hat{\mathbf{t}}_1$ direction, and v_2 is the quasi-S wave speed of waves polarized in the $\hat{\mathbf{t}}_2$ direction. The condition 2.16 perfectly absorbs waves impinging at a right angle to the boundary, but is less effective on waves that graze the boundary (Clayton & Engquist (44), 1977). It is valid for transversely isotropic media with a horizontal or vertical symmetry axis, but can be extended to more general anisotropy. A significant more efficient absorbing condition is, for example, the Perfectly Matched Layer (PML) (e.g., Bérenger (13), 1994; Komatitsch & Tromp (103), 2003; Chaljub et al. (32), 2007; Komatitsch & Martin (99), 2007; Martin et al. (134), 2008; Martin & Komatitsch (133), 2009).

At periods longer than about 150 s, the momentum equation 2.11 will involve additional terms, since the self-gravitation and rotation of the Earth should be taken into account modeling seismic wave propagation (Dahlen & Tromp (51), 1998). Dealing with these long period effects is beyond our scope.

2.1.3 Weak formulation of the problem

Many numerical methods, such as finite-difference and pseudospectral methods, are based upon a strong formulation of the problem; that is, they work directly with the equation of motion 2.11 and related boundary conditions written in differential form. Finite-element (FE) and spectral-element (SE) methods, instead, use a weak or integral formulation. This is obtained by first taking the dot product of the momentum equation 2.11 with an arbitrary test vector $\mathbf{w}$; then, integrating by parts over the volume Ω of

the model and imposing the boundary conditions 2.15 and 2.16, one obtains (e.g., Komatitsch et al. (110), 2005)

$$\int_\Omega \rho\, \mathbf{w} \cdot \partial_t^2 \mathbf{s}\, \mathrm{d}^3\mathbf{x} = -\int_\Omega \nabla \mathbf{w} : \mathbf{T}\, \mathrm{d}^3\mathbf{x} + \mathbf{M} : \nabla \mathbf{w}(\mathbf{x}_s)\, S(t) + \int_\Gamma \hat{\mathbf{n}} \cdot \mathbf{T} \cdot \mathbf{w}\, \mathrm{d}^2\mathbf{x}\,. \quad (2.17)$$

In a finite-element context, the term on the left hand side is called *mass matrix* and the first term on the right is the *stiffness matrix*.

The second term on the right hand side of 2.17 represents the source term $\int_\Omega \mathbf{f}\cdot\mathbf{w}\, \mathrm{d}^\mathbf{3}\mathbf{x}$ for a simple point source and it is obtained by explicitly integrating equation 2.12 using the properties of the Dirac delta distribution. In the case of a finite fault model it becomes, by integrating equation 2.13,

$$\int_{\Sigma_s} \mathbf{m}(\mathbf{x}_s, t) : \nabla \mathbf{w}(\mathbf{x}_s)\, \mathrm{d}^2\mathbf{x}_s\,. \quad (2.18)$$

The last term on the right in 2.17 arises for local and regional simulations. It involves the integral over the artificial absorbing model boundaries Γ of the traction $\hat{\mathbf{n}} \cdot \mathbf{T}$, that can be expressed in terms of the paraxial equation 2.16 (e.g., Komatitsch & Tromp (100), 1999).

Equation 2.17 is mathematically equivalent to the strong formulation 2.11, because it holds for any test vector $\mathbf{w}$. However, using a weak form of the equation of motion allows the traction-free surface condition 2.15 to be automatically satisfied during the integration by parts. In fact, the contour integral over the free surface $\partial\Omega$, that would have appeared in equation 2.17, simply vanishes. Condition 2.15 is, thereby, a natural condition of the problem, leading SEM to very accurately simulates surface wave propagation (e.g., Komatitsch & Vilotte (104), 1998; Komatitsch et al. (105), 1999).

Equation 2.17 is valid also for an attenuating medium, by representing the stress tensor as in 2.6 instead of 2.1. However, in this case one needs to solve equations 2.8 for the memory-variables, that significantly increases computer memory and time requirements (e.g., Komatitsch & Tromp (100), 1999; Komatitsch et al. (110), 2005).

In this study we won't take into account the case of an Earth model with fluid and solid coupled regions. Nevertheless, these complexities can be implemented in SEM, using for example a domain decomposition approach (Komatitsch & Tromp (101), 2002).

2.2 Spectral-element method (SEM)

2.2.1 Introduction

As discussed in section 1.1, accurately simulating seismic wavefields in complex media is of primary interest in seismology and a spectral-element method is very well suited to handle such a problem.

Next paragraphs outline, following for example Komatitsch & Tromp (100) (1999), Komatitsch et al. (110) (2005) and Tromp et al. (198) (2008), how the basic theory that governs local and regional seismic wave propagation is implemented in a spectral-element method in order to model seismic waveforms in detail. Moreover, in section 2.4, I will present the main features of the code SPECFEM3D version 2.0 (Peter et al. (160), 2011), that is based upon SEM on a local scale and has been used in this work.

As an introduction, in what follows I briefly summarize the key points of a SEM, that make this method particularly suitable for our needs and have led our choice. The following subsections will explain in more details all these aspects.

- SEM is based upon the weak formulation 2.17 of wave equation; thus, as discussed in section 2.1.3, the free-surface topography can be easily implemented and surface waves are as accurately simulated as body waves.

- Simulation volumes are discretized using a grid (*mesh*) of hexahedral elements (section 2.2.2). This mesh can honor any discontinuity in the model and can be fully unstructured (i.e., the number of elements that share a given point can vary and take any value), thus very complex geometries and any arbitrary shaped domain can be accommodated. Mechanical properties can vary inside each element, allowing fully heterogeneous media to be implemented. Moreover, a low numerical dispersion can be obtained with a small number of grid points per wavelength and the element size can be adjusted based upon the seismic frequency under consideration; this saves computing time and memory, and allows very high resolution simulations. Despite these favorable aspects, hexahedral meshing for realistic complex volumes is very challenging and more demanding than, for example, tetrahedral meshing (Shepherd & Johnson (178), 2008; Casarotti et al. (24), 2008; Staten et al. (181), 2010).

- As already stated in section 1.1, all the complexities of a 3D Earth model can be incorporated in simulating forward wavefields.

- SEM uses high-degree (between 4 and 10) Lagrange polynomials as basis functions to represent wavefields on the mesh elements (section 2.2.3). This ensures a very high spatial accuracy and an exponential decreasing of errors typical of spectral and pseudo-spectral methods.

- A Gauss-Lobatto-Legendre (GLL) quadrature is used to perform integrals in 2.17 at elemental level (section 2.2.3). This, together with the Lagrange interpolants defined at the GLL points of integration, give rise of an exactly diagonal mass matrix in the equation of motion 2.17. The SEM algorithm is drastically simplified and, as a consequence, it can be marched in time using a fully explicit integration scheme (section 2.2.4).

- The time scheme that solves the governing equation for the global system lends itself to be very efficiently implemented on parallel computers with distributed memory (section 2.2.4). This tremendously reduces the computational costs, making SEM suitable to be used for large, high-resolution simulations on very powerful machines (e.g., Komatitsch et al. (110), 2005).

2.2.2 Meshing

As in any FEM (e.g., Hughes (83), 1987), the first, crucial step in a SEM is to construct a high-quality mesh that discretizes the considered 3D domain.

Let us divide the model volume Ω into n_e non-overlapping, hexahedral elements Ω_e, e=1,..,n_e, such that $\Omega = \bigcup_e^{n_e} \Omega_e$. In SEM, unlike FEM, only hexahedral meshing is possible[1]. However, this leads to several benefits, such as optimized tensor products, a diagonal mass matrix and a smaller number of elements compared to tetrahedral meshing (e.g., Peter et al. (160), 2011).

The deformed volume elements Ω_e can be mapped to a reference cube, thus points $\mathbf{x}=(x, y, z)$ within each Ω_e are uniquely related to points $\xi=(\xi, \eta, \zeta)$, $-1 \leq \xi, \eta, \zeta \leq 1$,

[1]Note that also triangles may be used in SEMs (e.g., Komatitsch et al. (108), 2001), but this will introduce some complications we won't deal with.

in each cube by the invertible mapping

$$\mathbf{x}(\xi) = \sum_{a=1}^{n_a} \mathbf{x}_a N_a(\xi)\,. \tag{2.19}$$

The geometry of a hexahedral element is, therefore, defined in terms of n_a anchors or control points $\mathbf{x}_a$=$\mathbf{x}(\xi_a, \eta_a, \zeta_a)$, a=1,..,n_a, and n_a shape functions $N_a(\xi)$. The anchors are eight, if we consider only the corners of a volume element; they become 20 by adding the edge centers, and 27 by adding also the face centers and element center.

The shape functions $N_a(\xi)$ are usually products of Lagrange polynomials of degree one or two. The n+1 Lagrange polynomials of degree n are defined in terms of n+1 control points -1$\leq \xi_\alpha \leq$1, α=0,..,n, by

$$l_\alpha^n(\xi) = \frac{(\xi-\xi_0)...(\xi-\xi_{\alpha-1})(\xi-\xi_{\alpha+1})...(\xi-\xi_n)}{(\xi_\alpha-\xi_0)...(\xi_\alpha-\xi_{\alpha-1})(\xi_\alpha-\xi_{\alpha+1})...(\xi_\alpha-\xi_n)}\,. \tag{2.20}$$

As a consequence, at any control point ξ_β the Lagrange polynomials return either 0 or 1, that is

$$l_\alpha^n(\xi_\beta) = \delta_{\alpha\beta}\,, \tag{2.21}$$

where δ denotes the Kronecker delta. Hexahedral shape functions for an element described by 8 control nodes are triple products of degree 1 Lagrange polynomials, one for each direction in the reference cube; shape functions of 27-node hexahedral elements are triple products of degree 2 Lagrange polynomials.

An element of volume $\mathrm{d}^3\mathbf{x}$=$\mathrm{d}x\mathrm{d}y\mathrm{d}z$ within a hexahedral element Ω_e is related to an element of volume $\mathrm{d}^3\xi$=$\mathrm{d}\xi\mathrm{d}\eta\mathrm{d}\zeta$ in a reference cube by

$$\mathrm{d}^3\mathbf{x} = J_e \mathrm{d}^3\xi\,, \tag{2.22}$$

where the volumetric Jacobian of the mapping 2.19 is

$$J_e = \left|\frac{\partial(x,y,z)}{\partial(\xi,\eta,\zeta)}\right|\,. \tag{2.23}$$

The partial derivative matrix $\partial\mathbf{x}/\partial\xi$ in 2.23 is obtained by differentiating 2.19

$$\frac{\partial\mathbf{x}}{\partial\xi} = \sum_{a=1}^{n_a} \mathbf{x}_a \frac{\partial N_a}{\partial\xi}\,, \tag{2.24}$$

where the partial derivatives of the hexahedral shape functions can be calculated analytically in terms of degree 1 or 2 Lagrange polynomials and their derivatives.

The weak formulation 2.17 also involves a surface integral over the absorbing boundary Γ, that arises in the case of local and regional simulations. The mesh of hexahedra Ω_e should honor the major discontinuities in the model, as well as the absorbing edges. Thus, the surfaces are naturally divided into non-overlapping, quadrilateral elements Γ_b mapped to a reference square. Points $\mathbf{x}$ of any surface element are related to points (ξ, η) in the square in terms of anchors $\mathbf{x}_a$ and 2D shape functions $N_a(\xi, \eta)$, that is

$$\mathbf{x}(\xi, \eta) = \sum_{a=1}^{n_a} \mathbf{x}_a N_a(\xi, \eta)\,. \tag{2.25}$$

In this case, quadrilateral shape functions are products of two degree 1 or 2 Lagrange polynomials defined by 2.20 and 2.21. The normal $\hat{\mathbf{n}}$ to a boundary element Γ_b is defined as

$$\hat{\mathbf{n}} = \frac{1}{J_b} \frac{\partial \mathbf{x}}{\partial \xi} \times \frac{\partial \mathbf{x}}{\partial \eta}\,, \tag{2.26}$$

where the Jacobian J_b of the mapping 2.25,

$$J_b = \left|\left| \frac{\partial \mathbf{x}}{\partial \xi} \times \frac{\partial \mathbf{x}}{\partial \eta} \right|\right|\,, \tag{2.27}$$

defines the relation between an element of surface $\mathrm{d}^2\mathbf{x}$=dxdy in Γ_b and an element of surface $\mathrm{d}^2\xi$=dξdη in the reference square by $\mathrm{d}^2\mathbf{x}$= J_b $\mathrm{d}^2\xi$. The partial derivatives in 2.26 and 2.27 are calculated in terms of the partial derivatives $\partial_\xi N_a$ and $\partial_\eta N_a$ by analytically differentiating the mapping 2.25 (e.g., Komatitsch & Tromp (100), 1999).

The design of the mesh is subject to several constraints in order to perform stable and accurate calculations (e.g., Komatitsch et al. (110), 2005; Casarotti et al. (24), 2008):

- The mesh is required to be geometrically conforming, that is, the edges of each element should match up exactly with the edges of adjacent elements.
- The volume and surface elements should be constructed such that the Jacobians J_e (eq. 2.23) and J_b (eq. 2.27), respectively, should never vanish; this ensures that the mapping from the reference elements to the deformed ones is unique and invertible, i.e., from 2.19, for example, $\xi(\mathbf{x})$ is well-defined.
- The geometrical distortion of the elements shouldn't be too strong (e.g., Casarotti et al. (24), 2008), that is, only smooth local variations of the Jacobians are allowed within the mesh, in order to avoid inaccurate or unstable calculations.

- The number of grid points per shortest wavelength, i.e. the numerical resolution, defines the quality of the mesh (see also eq. 2.39). In order to maintain the same resolution everywhere in the model, the size of the elements should increase at increasing seismic wave speed.

- Finally, the mesh should satisfy the numerical stability condition imposed by the time scheme used to solve the governing equations (section 2.2.4).

2.2.3 Wavefield representation on elements and numerical integration

After meshing the model, the integrals over the volume Ω and the absorbing surface Γ in equation 2.17 are subdivided into smaller integrals over the volume elements Ω_e and the surface elements Γ_b, respectively. Thus, we need to represent at the elemental level the involved functions, such as the displacement $\mathbf{s}$ and the test vector $\mathbf{w}$, and to define the corresponding integration rule.

As seen in the previous section, in a SEM (as in a FEM) the geometry of the mesh elements is expressed in terms of low-degree (1 or 2) Lagrange polynomials. In a traditional FEM these low-degree polynomials are also used to represent the functions on an element. On the contrary, SEMs use higher-degree Lagrange interpolants as basis functions to discretize the wavefields. This is one of the most relevant differences between SEMs and FEMs. Usually, Lagrange polynomials of degree between 4 and 10 are considered, which guarantees a high spatial accuracy but still affordable computational costs (e.g., Seriani & Priolo (177), 1994; Komatitsch & Vilotte (104), 1998; Komatitsch et al. (110), 2005).

In a SEM, the control points ξ_α, α=0,..,n, needed in the definition 2.20 of the n+1 Lagrange polynomials of degree n, are chosen to be the n+1 Gauss-Lobatto-Legengre (GLL) points, that are calculated numerically as roots of the equation (Canuto et al. (21), 1988)

$$(1-\xi^2)P_n'(\xi) = 0\,, \tag{2.28}$$

where P_n' is the derivative of the Legendre polynomial of degree n. Values -1 and 1 are always roots of 2.28, thus in a SEM some GLL points always lie exactly on the boundaries of the elements. For Lagrange polynomials of degree n, each 3D spectral element contains $(n+1)^3$ non-evenly spaced GLL points, and each 2D face contains $(n+1)^2$ GLL points (Fig. 2.1).

As shown in what follows, using high-degree Lagrange polynomials, defined at the GLL points, to interpolate wavefield functions is a crucial choice. In fact, these interpolants, in conjunction with a particular integration rule, result in an exactly diagonal mass matrix. This, in turn, leads to a simple explicit time scheme extremely suitable for parallel implementation (section 2.2.4).

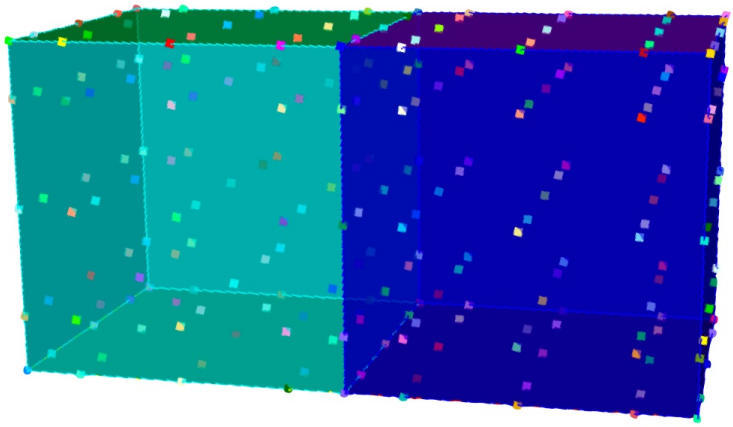


Figure 2.1: GLL point distribution - For Lagrange polynomials of degree n=4, each edge of a hexahedral element contains 5 non-evenly spaced GLL points, thus, each 2D face contains $(n+1)^2$=25 GLL points and each 3D element contains $(n+1)^3$=125 GLL points.

Functions on an element

On a volume element Ω_e, a function f may be represented in terms of Lagrange polynomials with degree n (eq. 2.20) and GLL control points as

$$f(\mathbf{x}(\xi,\eta,\zeta)) \approx \sum_{\alpha,\beta,\gamma=0}^{n} f^{\alpha\beta\gamma} l_\alpha(\xi) l_\beta(\eta) l_\gamma(\zeta)\,. \tag{2.29}$$

Due to definition 2.21 of the Lagrange polynomials, the coefficients $f^{\alpha\beta\gamma} = f(\mathbf{x}(\xi_\alpha,\eta_\beta,\zeta_\gamma))$ are the values of the function at the GLL interpolation points $\mathbf{x}(\xi_\alpha,\eta_\beta,\zeta_\gamma)$. The polynomial degree n is omitted in 2.29.

Analogously, on a surface element Γ_b, f is expressed as

$$f(\mathbf{x}(\xi,\eta)) \approx \sum_{\alpha,\beta=0}^{n} f^{\alpha\beta} l_\alpha(\xi) l_\beta(\eta)\,, \tag{2.30}$$

where $f^{\alpha\beta} = f(\mathbf{x}(\xi_\alpha, \eta_\beta))$.

Equation 2.17 involves gradients of vectors $\mathbf{s}$ and $\mathbf{w}$. Based upon relation 2.29, the gradient of a function f, evaluated at the GLL point $\mathbf{x}(\xi_{\alpha'}, \eta_{\beta'}, \zeta_{\gamma'})$, may be written as

$$\nabla f(\mathbf{x}(\xi_{\alpha'}, \eta_{\beta'}, \zeta_{\gamma'})) \approx \sum_{i=1}^{3} \hat{\mathbf{x}}_i \left[(\partial_i \xi)^{\alpha'\beta'\gamma'} \sum_{\alpha=0}^{n} f^{\alpha\beta'\gamma'} l'_\alpha(\xi_{\alpha'}) + (\partial_i \eta)^{\alpha'\beta'\gamma'} \sum_{\beta=0}^{n} f^{\alpha'\beta\gamma'} l'_\beta(\eta_{\beta'}) + (\partial_i \zeta)^{\alpha'\beta'\gamma'} \sum_{\gamma=0}^{n} f^{\alpha'\beta'\gamma} l'_\gamma(\zeta_{\gamma'}) \right], \tag{2.31}$$

where $x_1{=}x$, $x_2{=}y$, $x_3{=}z$, $\partial_i{=}\partial_{x_i}$, and a prime denotes derivatives of the Lagrange polynomials. The matrix $\partial\xi/\partial\mathbf{x}$ is calculated as the inverse of the Jacobian matrix $\partial\mathbf{x}/\partial\xi$, that exists because it is one of the constrain of the mesh design (section 2.2.2).

Integration over elements

In a classical FEM integrals in 2.17 are numerically evaluated at the elemental level using a Gauss quadrature. On the contrary, SEM uses a Gauss-Lobatto-Legendre (GLL) integration rule, where points of integration are exactly the GLL interpolants, hence leading to the exactly diagonal mass matrix.

Based upon GLL quadrature, integrals over the volume elements Ω_e of a function f are expressed as

$$\int_{\Omega_e} f(\mathbf{x}) \mathrm{d}^3\mathbf{x} = \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} f(\mathbf{x}(\xi,\eta,\zeta)) J_e(\xi,\eta,\zeta) \mathrm{d}\xi\, \mathrm{d}\eta\, \mathrm{d}\zeta \approx \sum_{\alpha,\beta,\gamma=0}^{n} \omega_\alpha \omega_\beta \omega_\gamma f^{\alpha\beta\gamma} J_e^{\alpha\beta\gamma}, \tag{2.32}$$

where $J_e^{\alpha\beta\gamma} = J_e(\xi_\alpha, \eta_\beta, \zeta_\gamma)$, and $\omega_\alpha > 0$, $\alpha{=}0,..,n$, denote the $n{+}1$ weights of the GLL integration points, that are numerically estimated (Canuto et al. (21), 1988).

Similarly, the integrals over the surface element Γ_b of the absorbing boundary may be written as

$$\int_{\Gamma_b} f(\mathbf{x}) \mathrm{d}^2\mathbf{x} = \int_{-1}^{1}\int_{-1}^{1} f(\mathbf{x}(\xi,\eta)) J_b(\xi,\eta) \mathrm{d}\xi\, \mathrm{d}\eta \approx \sum_{\alpha,\beta=0}^{n} \omega_\alpha \omega_\beta f^{\alpha\beta} J_b^{\alpha\beta}, \tag{2.33}$$

where $J_b^{\alpha\beta} = J_b(\xi_\alpha, \eta_\beta)$.

Other authors (e.g., Priolo et al. (166), 1994; Seriani (176), 1988) use an implementation of SEM based upon Chebyshev polynomials. The Gauss-Lobatto-Chebyshev (GLC) integration rule is exact for the chosen polynomial basis; the GLL quadrature,

instead, is exact only for polynomials of degree $2n$-1, thus, never exact in SEM[1]. Moreover, the GLC points and weights can be analytically determined. However, in this case the advantage of a diagonal mass matrix is lost (e.g., Komatitsch et al. (110), 2005).

2.2.4 Global system time-marching

In the previous section we have defined the basis functions for wavefield interpolation and the numerical integration rule used by SEM. At this point, the equation of motion 2.17 in weak formulation may be discretized at the elemental level. In order to approximate the term on the left hand side, i.e., the *mass matrix*, we firstly expand the displacement vector $\mathbf{s}$ and the test vector $\mathbf{w}$ using the polynomial representation in 2.29

$$\mathbf{s}(\mathbf{x}(\xi,\eta,\zeta),t) \approx \sum_{j=1}^{3} \hat{\mathbf{x}}_j \sum_{\sigma,\tau,\nu=0}^{n} s_j^{\sigma,\tau,\nu}(t) l_\sigma(\xi) l_\tau(\eta) l_\nu(\zeta)\,, \tag{2.34}$$

$$\mathbf{w}(\mathbf{x}(\xi,\eta,\zeta)) = \sum_{i=1}^{3} \hat{\mathbf{x}}_i \sum_{\alpha,\beta,\gamma=0}^{n} w_i^{\alpha\beta\gamma} l_\alpha(\xi) l_\beta(\eta) l_\gamma(\zeta)\,. \tag{2.35}$$

Since both the displacement and the test vectors are expressed in terms of the same basis functions, the SEM is a so-called Galerkin method. Next, using the GLL quadrature rule in 2.32 and the expressions of $\mathbf{s}$ and $\mathbf{w}$ in 2.34 and 2.35, the *mass matrix* on hexahedral elements Ω_e may be written as (e.g., Komatitsch & Tromp (100), 1999)

$$\begin{aligned} \int_{\Omega_e} \rho\, \mathbf{w}\cdot\partial_t^2\mathbf{s}\,\mathrm{d}^3\mathbf{x} &= \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} \rho(\mathbf{x}(\xi))\mathbf{w}(\mathbf{x}(\xi))\cdot\partial_t^2\mathbf{s}(\mathbf{x}(\xi),t) J_e(\xi)\mathrm{d}^3\xi \\ &\approx \sum_{\alpha,\beta,\gamma} \omega_\alpha\omega_\beta\omega_\gamma J_e^{\alpha\beta\gamma}\rho^{\alpha\beta\gamma}\sum_{i=1}^{3} w_i^{\alpha\beta\gamma}\ddot{s}_i^{\alpha\beta\gamma}(t)\,, \end{aligned} \tag{2.36}$$

where $\rho^{\alpha\beta\gamma}$=$\rho(\mathbf{x}(\xi_\alpha,\eta_\beta,\zeta_\gamma))$, and a dot indicates differentiation with respect to time. The density ρ may be different at each gridpoint, thus SEM can implement fully heterogeneous media. It is worth noting that the weak equation 2.17 holds for any test vector $\mathbf{w}$. Therefore, one can independently set factors of $w_1^{\alpha\beta\gamma}$, $w_2^{\alpha\beta\gamma}$ and $w_3^{\alpha\beta\gamma}$ equal to zero, which results in independent equations for each component of acceleration $\ddot{s}_i^{\alpha\beta\gamma}$(t) at grid point $(\xi_\alpha,\eta_\beta,\zeta_\gamma)$. As a consequence of 2.36, the value of acceleration at each point of a given element, $\ddot{s}_i^{\alpha\beta\gamma}$(t), is simply multiplied by the factor $(\omega_\alpha\omega_\beta\omega_\gamma J_e^{\alpha\beta\gamma}\rho^{\alpha\beta\gamma})$, that

[1] In SEM, integrands are of degree $2n$, since they involve products of two polynomials of degree n, the displacement and the test function.

is, the elemental *mass matrix* is exactly diagonal. This is one of the key points of SEM, and the main reason of the choice of Lagrange interpolation at GLL points in conjunction with GLL numerical quadrature (section 2.2.3).

Similarly to equation 2.36, the other terms in 2.17, i.e., the *stiffness matrix*, the source term and the integral over the absorbing model boundary, may be written in a discrete form. Explicit expressions of these terms at the elemental level, by using expansions in 2.34 and 2.35, the GLL integration rule and the constitutive relationships in section 2.1.1, are thoroughly reported e.g. in Komatitsch & Tromp (100) (1999).

As previously shown, in a SEM the volume is subdivided into hexahedral elements, and the functions are interpolated at the GLL integration points within each element. Neighboring hexahedra share points on their sides, faces and corners. Thus, one needs to determine a mapping between the gridpoints that define an element, the *local points*, and the collective points of the system, the *global points*, many of which are shared amongst different elements. This can be accomplished by using finite-element routines (e.g., Komatitsch et al. (110), 2005).

The next step requires to sum up, at each point of the global mesh, the contribution from all the elements that share a common global gridpoint. This stage is called the *assembly* of the system and involves exchange of information between neighboring elements.

One of the main advantages of SEM is that it can be easily implemented on parallel computers, that drastically reduces the computational costs. To this purpose, the mesh is subdivided into slices of elements, each of which is assigned to a different processor. The distribution is accomplished by a partitioner (such as SCOTCH, Chevalier & Pellegrini (37), 2008) and is based upon a *load-balancing* logic (e.g., Peter et al. (160), 2011), i.e., the mesh is partitioned such that each processor involved in the calculation performs roughly the same number of operations per time step. The standard way of programming such parallel machines with distributed memory is to use a message passing methodology, that is based upon the Message Passing Interface (MPI) (e.g., Gropp et al. (80), 1994)[1]. In this context, the assembling of the system requires that, at each time step, the values of the forces, computed separately on each element, are summed at common gridpoints shared by elements of adjacent slices. Thus, it involves communication between distinct CPUs, that is a very expensive process on parallel computers.

[1]More details on parallel implementation of SEM may be found in, e.g., Peter et al. (160) (2011).

The use of a particular time scheme can help in reducing the computational load, as shown in what follows.

After assembling, the differential equation that governs the global system may be deduced. Let us define the global displacement vector $\boldsymbol{U}$; its components are the displacement vector at all the global grid points, classically referred to as the *global degrees of freedom* of the system. The corresponding global test vector is denoted by $\boldsymbol{W}$. Moreover, from the elemental form of each term of equation 2.17 (e.g. 2.36), one can construct, respectively, the *global* mass matrix $\boldsymbol{M}$, the *global* stiffness matrix $\boldsymbol{K}$, the source term $\boldsymbol{F}$ for the global system and the *global* absorbing boundary matrix $\boldsymbol{C}$ (e.g., Komatitsch & Tromp (100), 1999; Komatitsch & Vilotte (104), 1998). All these terms combine into the second-order ordinary differential equation that governs the time evolution of the global system:

$$\boldsymbol{M}\,\ddot{\boldsymbol{U}} = -\boldsymbol{K}\,\boldsymbol{U} + \boldsymbol{F} - \boldsymbol{C}\,\dot{\boldsymbol{U}}\,. \tag{2.37}$$

Since the elemental mass matrix is diagonal, after the assembly of the system, also the global mass matrix $\boldsymbol{M}$ results to be diagonal.

As mentioned above, one can take full advantage of a diagonal global mass matrix by using a fully explicit second-order finite-difference time scheme to march the differential equation 2.37. This scheme is a particular case of a more general Newmark one (Hughes (83), 1987) and is conditionally stable, i.e., for a given mesh and a given model there exists an upper limit on the time step above which the calculations become unstable. The Courant number of the integration scheme may be defined as $C = \Delta t(v/\Delta h)_{max}$, where Δt is the time step, and $(v/\Delta h)_{max}$ is the maximum ratio between the P-wave speed and the grid spacing Δh. The Courant stability condition (Courant et al. (47), 1928) requires that $C \leq C_{max}$, where C_{max} is the maximum Courant number. This means that the time step should not be chosen higher than an upper limit:

$$\Delta t \leq C_{max} \left(\frac{\Delta h}{v} \right)_{min} . \tag{2.38}$$

Based upon a heuristic rule, C_{max} is roughly equal to 0.3 or 0.4 for irregular meshes and heterogeneous media (e.g., Komatitsch et al. (110), 2005). Thanks to this explicit time scheme, the SEM algorithm implemented in a parallel machine mostly consists of small local calculations, performed by each processor on the elements of the slice it carries. Instead, the most expensive phase, i.e., the communication between CPUs

required at each time step, represents only a small fraction of the computation time, as desirable.

For anelastic media, the strong form of the memory-variable equation 2.8 is solved separately for $\mathbf{R}_l$ by using a modified second-order Runge-Kutta scheme in time (Carcione (23), 1994; Komatitsch et al. (110), 2005).

To conclude, concerning the accuracy of the method, in a SEM this is controlled by both the grid spacing Δh in the mesh, and by the degree n of the Lagrange polynomials used to represent functions on elements. On the contrary, in a classical FEM, functions are expressed in terms of low-degree polynomials (section 2.2.3), thus only the element size is relevant to adjust the accuracy. In the SEM, the following heuristic rule can be used to choose the average grid spacing Δh for a given polynomial degree n (e.g., Komatitsch et al. (110), 2005):

$$\Delta h = \lambda_{min} \frac{n+1}{f(n)} = T_0 \, v_{min} \frac{n+1}{f(n)} \,, \tag{2.39}$$

where $f(n)$ describes the average number of grid points per minimum wavelength λ_{min} in an element and depends on n; T_0 is the minimum period that one seeks to resolve, and v_{min} is the minimum wave speed in the element[1]. Thus, in order to obtain accurate results, for n=4, Δh should be chosen such that $f(n)$ is roughly equal to 5.

Based upon equation 2.39, the numerical stability condition 2.38, that defines the choice of the time step in the time scheme for SEM, may be rewritten as

$$\Delta t \leq C_{max} \, T_0 \, \frac{v_{min}}{v_{max}} \, \frac{n+1}{f(n)} \,. \tag{2.40}$$

2.3 Adjoint methods

2.3.1 General formulation of the inverse problem

As introduced in section 1.1, the objective is to minimize a measure of the misfit between a set of observed data and a corresponding set of synthetics computed for a model $\mathbf{m}$. Several choices are possible for the *misfit function* $F(\mathbf{m})$, depending on the chosen measure of the misfit between the two time series, for example a least-squares

[1]This is the S-wave speed for elements inside the model and the surface-wave speed for elements at the free surface (Komatitsch et al. (110), 2005).

waveform difference (section 2.3.2) or a travel time difference (section 2.3.3). The model vector $\mathbf{m}$ contains, for example, the values of structural and source parameters and we seek to determine a model correction $\delta\mathbf{m}$ which leads to a minimum of the misfit function F (Nolet (147), 1987; Tarantola (193), 2005). The quadratic approximation of $F(\mathbf{m}+\delta\mathbf{m})$ may be written (e.g., Tape et al. (187), 2007) [1]:

$$F(\mathbf{m}+\delta\mathbf{m}) \approx F(\mathbf{m}) + \mathbf{g}(\mathbf{m})^T \delta\mathbf{m} + \frac{\mathbf{1}}{\mathbf{2}} \delta\mathbf{m}^T \mathbf{H}(\mathbf{m}) \delta\mathbf{m}\,, \tag{2.41}$$

where the gradient vector, evaluated at $\mathbf{m}$, i.e. the first derivative of the misfit function or Fréchet derivative, is

$$\mathbf{g}(\mathbf{m}) = \left.\frac{\partial F}{\partial \mathbf{m}}\right|_{\mathbf{m}}\,, \tag{2.42}$$

and the Hessian matrix, evaluated at $\mathbf{m}$, is defined in terms of the second derivatives of the misfit function by

$$\mathbf{H}(\mathbf{m}) = \left.\frac{\partial^{\mathbf{2}} F}{\partial \mathbf{m} \partial \mathbf{m}}\right|_{\mathbf{m}}\,. \tag{2.43}$$

The gradient of equation 2.41 with respect to $\delta\mathbf{m}$ is

$$\mathbf{g}(\mathbf{m}+\delta\mathbf{m}) \approx \mathbf{g}(\mathbf{m}) + \mathbf{H}(\mathbf{m})\delta\mathbf{m}\,, \tag{2.44}$$

which can be set equal to zero to obtain the model correction $\delta\mathbf{m}$ that minimizes 2.41:

$$\mathbf{H}(\mathbf{m})\delta\mathbf{m} = -\mathbf{g}(\mathbf{m}), \tag{2.45}$$

$$\delta\mathbf{m} = -\mathbf{H}^{-\mathbf{1}}(\mathbf{m})\,\mathbf{g}(\mathbf{m})\,. \tag{2.46}$$

An updated model $\mathbf{m}+\delta\mathbf{m}$ can be obtained with or without the Hessian $\mathbf{H}$. As introduced in section 1.1, approaches to solve the inverse problem that can access to both gradient and Hessian are called *Newton methods*; this is generally the case of classical traveltime tomography, that is based upon equations 2.45 - 2.46. Considering complex 3D heterogeneous models, only the calculation of the gradient in 2.42 is feasible; then one needs to solve the inverse problem using iterative gradient-based approaches, e.g. the steepest descent or conjugate gradient methods, as in adjoint tomography (e.g., Tape et al. (187, 189), 2007, 2010).

[1] It comes from the Taylor series expansion of a function $f(x_0+\delta x)$ about a point x_0 by substituting $x_0 \rightarrow \mathbf{m}$ and $\delta x \rightarrow \delta\mathbf{m}$

In this work, we are interested in iteratively improve an initial 3D velocity model for central Italy lithosphere by using an adjoint approach and a steepest descent algorithm to minimize traveltime differences between observed and synthetic waveforms. In order to introduce the general basic formulation of the adjoint methods, I will start by considering a waveform tomography (section 2.3.2). Then, in section 2.3.3, I will discuss traveltime tomography based upon the adjoint approach, highlighting connections and differences with respect to a classical traveltime tomography based upon finite-frequency kernels. Other choices for the misfit function are discussed, e.g., in Tromp et al. (197) (2005), leading to different tomographic approaches.

2.3.2 Waveform tomography and general adjoint equations

Following Tromp et al. (197) (2005), in waveform tomography one can choose to minimize a *waveform misfit function* defined in terms of the least-squares difference between three-component waveform data $\mathbf{d}(\mathbf{x}_r,\mathrm{t})$ and the corresponding synthetics $\mathbf{s}(\mathbf{x}_r, t, \mathbf{m})$ for a given model $\mathbf{m}$ (Nolet (147), 1987):

$$F(\mathbf{m}) = \frac{1}{2}\sum_{r=1}^{N}\int_0^T ||\mathbf{s}(\mathbf{x}_r, t, \mathbf{m}) - \mathbf{d}(\mathbf{x}_r, t)||^2 \,\mathrm{d}t\,, \tag{2.47}$$

where $\mathbf{x}_r$, r=1,...,N, are the coordinates of the N receivers, and the model vector $\mathbf{m}$ consists of M components. The data $\mathbf{d}$ and synthetics $\mathbf{s}$ are assumed to be processed, i.e. they have been already windowed and filtered on the time window [0, T]. The misfit function 2.47 may include the uncertainties associated to the differences between data and synthetics, that give rise to a data covariance matrix $\boldsymbol{C}_D$ (refer to section 2.3.4). We do not include this weighting for sake of simplicity. In addition, norms other than L^2 may be introduced (Crase et al. (48), 1990).

The variation of the waveform misfit function 2.47 due to a model perturbation $\delta\mathbf{m}$ is given by

$$\delta F(\mathbf{m}) = \sum_{r=1}^{N}\int_0^T [\mathbf{s}(\mathbf{x}_r, t, \mathbf{m}) - \mathbf{d}(\mathbf{x}_r, t)] \cdot \delta\mathbf{s}(\mathbf{x}_r, t, \mathbf{m})\,\mathrm{d}t\,, \tag{2.48}$$

where $\delta\mathbf{s}$ is the perturbation in the displacement field $\mathbf{s}$ due to $\delta\mathbf{m}$. A generic anisotropic model may be described in terms of the density ρ and the fourth-order elastic tensor $\mathbf{c}$ with components c_{jklm}; the corresponding perturbations are, respectively, $\delta\rho$ and

δc_{jklm}. Using index notation, the i-th component of the perturbed displacement $\delta\mathbf{s}$, based upon the Born approximation (Hudson (82), 1977; Wu & Aki (206), 1985), may be written as (e.g., Tromp et al. (197), 2005):

$$\begin{aligned}\delta s_i(\mathbf{x},t) = -\int_0^t \int_V \big[\delta\rho(\mathbf{x}')\, G_{ij}(\mathbf{x},\mathbf{x}';t-t')\, \partial_{t'}^2 s_j(\mathbf{x}',t') \\ + \delta c_{jklm}(\mathbf{x}')\, \partial'_k G_{ij}(\mathbf{x},\mathbf{x}';t-t')\, \partial'_l s_m(\mathbf{x}',t')\big]\, \mathrm{d}^3\mathbf{x}'\, \mathrm{d}t'\,,\end{aligned} \tag{2.49}$$

where V is the model volume, G_{ij} are the components of the Green's tensor and from here on we neglect the dependence on $\mathbf{m}$. Substituting 2.49 in 2.48 we obtain

$$\begin{aligned}\delta F = -\sum_{r=1}^{N} \int_0^T [s_i(\mathbf{x}_r,t) - d_i(\mathbf{x}_r,t)] \int_0^t \int_V \big[\delta\rho(\mathbf{x}')\, G_{ij}(\mathbf{x}_r,\mathbf{x}';t-t')\, \partial_{t'}^2 s_j(\mathbf{x}',t') \\ + \delta c_{jklm}(\mathbf{x}')\, \partial'_k G_{ij}(\mathbf{x}_r,\mathbf{x}';t-t')\, \partial'_l s_m(\mathbf{x}',t')\big]\, \mathrm{d}^3\mathbf{x}'\mathrm{d}t'\mathrm{d}t\,.\end{aligned} \tag{2.50}$$

One can introduce the field

$$\Phi_k(\mathbf{x}',t') = \sum_{r=1}^{N} \int_{t'}^{T} G_{ik}(\mathbf{x}_r,\mathbf{x}';t-t')\, [s_i(\mathbf{x}_r,t) - d_i(\mathbf{x}_r,t)]\, \mathrm{d}t\,, \tag{2.51}$$

that, using the reciprocity of the Green's tensor (Aki & Richards (3), 1980; Dahlen & Tromp (51), 1998)

$$G_{ik}(\mathbf{x}_r,\mathbf{x}';t-t') = G_{ki}(\mathbf{x}',\mathbf{x}_r;t-t')\,, \tag{2.52}$$

and reversing time by substituting $t \to T-t$, may be rewritten as

$$\Phi_k(\mathbf{x}',t') = \sum_{r=1}^{N} \int_0^{T-t'} G_{ki}(\mathbf{x}',\mathbf{x}_r;T-t-t')\, [s_i(\mathbf{x}_r,T-t) - d_i(\mathbf{x}_r,T-t)]\, \mathrm{d}t\,. \tag{2.53}$$

At this point, one can define the function

$$f_i^{\dagger}(\mathbf{x},t) = \sum_{r=1}^{N} [s_i(\mathbf{x}_r,T-t) - d_i(\mathbf{x}_r,T-t)]\, \delta(\mathbf{x}-\mathbf{x}_r)\,, \tag{2.54}$$

that can be thought as the source of a so-called adjoint field in waveform tomography, i.e. it is the *waveform adjoint source*, constructed by considering the time-reversed differences between data and synthetics located at each receiver as N simultaneous point sources. Thus, using 2.54, the field in 2.53 may be written as

$$\Phi_k(\mathbf{x}',t') = \int_0^{T-t'} \int_V G_{ki}(\mathbf{x}',\mathbf{x};T-t-t') f_i^{\dagger}(\mathbf{x},t)\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}t\,, \tag{2.55}$$

and, by using the standard Green's function approach, one can finally introduce the *waveform adjoint field* $\mathbf{s}^\dagger$ generated by the waveform adjoint source 2.54

$$s_k^\dagger(\mathbf{x}',t') = \int_0^{t'} \int_V G_{ki}(\mathbf{x}',\mathbf{x};t'-t) f_i^\dagger(\mathbf{x},t)\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}t\,, \tag{2.56}$$

that is related to 2.55 by

$$\Phi_k(\mathbf{x}',t') = s_k^\dagger(\mathbf{x}',T-t')\,. \tag{2.57}$$

The adjoint wavefield $\mathbf{s}^\dagger$ is, thereby, generated based upon time-reversed signals at the stations and, in general, it must satisfy all the same equations as the regular forward displacement field $\mathbf{s}$ (section 2.1.2), except for the fact that adjoint quantities are now involved in the equations [1]. In particular, $\mathbf{s}^\dagger$ is obtained by solving the adjoint seismic-wave equation

$$\rho\, \partial_t^2 \mathbf{s}^\dagger = \nabla\cdot\mathbf{T}^\dagger + \mathbf{f}^\dagger\,, \tag{2.58}$$

where the adjoint stress is given by (compared to 2.1)

$$\mathbf{T}^\dagger = \mathbf{c} : \nabla\mathbf{s}^\dagger\,, \tag{2.59}$$

or for the anelastic case (compared to 2.6, 2.8, 2.10)

$$\mathbf{T}^\dagger = \mathbf{c}_U : \nabla\mathbf{s}^\dagger - \sum_{l=1}^{L} \mathbf{R}_l^\dagger\,, \tag{2.60}$$

with

$$\partial_t \mathbf{R}_l^\dagger = -(\mathbf{R}_l^\dagger - \delta\mu_l \mathbf{D}^\dagger)/\tau_l^\sigma\,, \tag{2.61}$$

$$\mathbf{D}^\dagger = \frac{1}{2}[\nabla\mathbf{s}^\dagger + (\nabla\mathbf{s}^\dagger)^T] - \frac{1}{3}(\nabla\cdot\mathbf{s}^\dagger)\,\mathbf{I}\,. \tag{2.62}$$

As shown above, the main difference, comparing 2.11 and 2.58, is that the regular field $\mathbf{s}$ is determined by the source term $\mathbf{f}$, that represents the earthquake (equations 2.12 or 2.13), whereas the adjoint field $\mathbf{s}^\dagger$ is generated by the adjoint source $\mathbf{f}^\dagger$, that, in waveform tomography, is given by (from 2.54)

$$\mathbf{f}^\dagger(\mathbf{x},t) = \sum_{r=1}^{N} [\mathbf{s}(\mathbf{x}_r, T-t) - \mathbf{d}(\mathbf{x}_r, T-t)]\, \delta(\mathbf{x}-\mathbf{x}_r)\,. \tag{2.63}$$

[1] The adjoint equations and the following kernel formulations could have been also explicitly derived by using a Lagrange multiplier approach as in Liu & Tromp (118), (2006)

Thus, from 2.63, in waveform tomography the simultaneously back-propagated signals are time-reversed differences between observed and synthetic waveforms at stations. In other tomographic approaches, different signals are sent back (see section 2.3.3 for adjoint traveltime tomography), thus $\mathbf{f}^\dagger$ has different forms other than 2.63, however 2.58 - 2.62 must hold in general. Finally, the set of adjoint wave equations needs to satisfy the initial conditions

$$\mathbf{s}^\dagger(\mathbf{x},0) = \mathbf{0}, \qquad \partial_t \mathbf{s}^\dagger(\mathbf{x},0) = \mathbf{0}\,, \tag{2.64}$$

the boundary condition at the free surface

$$\hat{\mathbf{n}} \cdot \mathbf{T}^\dagger = \mathbf{0}\,, \tag{2.65}$$

and, at solid-solid boundaries, both adjoint traction $\hat{\mathbf{n}} \cdot \mathbf{T}^\dagger$ and adjoint displacement $\mathbf{s}^\dagger$ must be continuous, whereas at fluid-solid boundaries adjoint traction $\hat{\mathbf{n}} \cdot \mathbf{T}^\dagger$ and normal component of adjoint displacement $\hat{\mathbf{n}} \cdot \mathbf{s}^\dagger$ must be continuous.

With the definition of the adjoint wavefield 2.56, the variation of the misfit function 2.50 can be rewritten as (following e.g. Tromp et al. (197), 2005)

$$\delta F = \int_V \left(K_\rho(\mathbf{x})\, \delta \ln \rho(\mathbf{x}) + K_{c_{jklm}}(\mathbf{x})\, \delta c_{jklm}(\mathbf{x}) \right) \mathrm{d}^3\mathbf{x}\,, \tag{2.66}$$

where δc_{jklm} denotes perturbations in the elastic tensor and $\delta \ln \rho$ is more precisely $\delta \ln(\rho/\rho_0) = \delta\rho/\rho$, i.e., the relative perturbation in density with ρ_0 a scaling parameter. Equation 2.66 shows a key point, i.e., in an adjoint tomographic inversion (as in Talagrand & Courtier (185) (1987) for general formulation of adjoint methods), the variation of the misfit function may be rewritten as a volume integral, which involves the model perturbations and 3D kernels, that are the Fréchet derivatives of F with respect to model parameters (see also section 2.3.3).

For the density and the elastic tensor the *waveform misfit kernels* are, respectively:

$$K_\rho = -\rho \int_0^T \mathbf{s}^\dagger(T-t) \cdot \partial_t^2 \mathbf{s}(t)\, \mathrm{d}t\,, \tag{2.67}$$

$$K_{c_{jklm}} = -\int_0^T \epsilon^\dagger_{jk}(T-t)\, \epsilon_{lm}(t)\, \mathrm{d}t\,, \tag{2.68}$$

where ϵ_{lm} and $\epsilon^\dagger_{jk}$ denote elements of the strain and adjoint strain tensors. Here and in what follows we have neglected the dependence on $\mathbf{x}$ to avoid clutter.

For an isotropic model, based upon relationship 2.2, equation 2.66 becomes

$$\delta F = \int_V \left(K_\rho \, \delta \ln \rho + K_\mu \, \delta \ln \mu + K_\kappa \, \delta \ln \kappa \right) \, \mathrm{d}^3\mathbf{x} \, , \tag{2.69}$$

where the isotropic misfit kernels K_μ and K_κ are the Fréchet derivatives with respect to relative shear and bulk modulus perturbations, respectively:

$$K_\mu = \, -2\mu \int_0^T \mathbf{D}^\dagger(T-t) : \mathbf{D}(t) \, \mathrm{d}t \, , \tag{2.70}$$

$$K_\kappa = \, -\kappa \int_0^T \left[\nabla \cdot \mathbf{s}^\dagger(T-t) \right] \left[\nabla \cdot \mathbf{s}(t) \right] \, \mathrm{d}t \, , \tag{2.71}$$

with $\mathbf{D}$ given by 2.10 and $\mathbf{D}^\dagger$ given by 2.62.

Alternatively, in the isotropic case, one can use the relationship between (κ, μ, ρ) and the compressional and shear wave speeds v_p and v_s, to express δF as a function of the compressional and shear wave speed kernels K_α and K_β, respectively (e.g., Tromp et al. (197), 2005)

$$\delta F = \int_V \left(K'_\rho \, \delta \ln \rho + K_\alpha \, \delta \ln v_p + K_\beta \, \delta \ln v_s \right) \, \mathrm{d}^3\mathbf{x} \, , \tag{2.72}$$

where

$$K'_\rho = K_\kappa + K_\mu + K_\rho \, , \qquad K_\alpha = 2 \left(\frac{\kappa + \frac{4}{3}\mu}{\kappa} \right) K_\kappa \, , \qquad K_\beta = 2 \left(K_\mu - \frac{4}{3} \frac{\mu}{\kappa} \, K_\kappa \right) \, . \tag{2.73}$$

Another possible parametrization for isotropic models is to use bulk sound wave speed $v_b = \sqrt{\kappa/\rho}$, shear wave speed v_s and density ρ (Tarantola (191), 1987), since v_b and v_s are independent combinations of the bulk and shear moduli κ and μ (e.g., Tape et al. (189), 2010). As shown in section 4.2.1, in our tomographic inversion we will use v_p and v_s; in future developments of this work we could try to use v_b.

It is worth noting that all the kernels in 2.66, 2.69 and 2.72 are obtained based upon the interaction between the regular wavefield $\mathbf{s}$ and the adjoint wavefield $\mathbf{s}^\dagger$, as introduced in section 1.1. Thus, for each event, the so-called *event kernels* are directly constructed by performing one simulation for the regular field, which propagates forward in time from source to stations, and one simulation for the adjoint wavefield, which propagates in reverse time from stations to source and is generated by simultaneous adjoint sources at receivers. So-called *misfit kernels* are obtained by simply summing

the event kernels for all the considered events and they represent the overall sensitivity to the model perturbations, highlighting where the current model needs to be modified to minimize the misfit between data and synthetics.

2.3.3 Traveltime adjoint tomography

In chapter 4 we will perform a traveltime tomography for central Italy using the SEM for wavefield simulations and the adjoint method for the inversion. Thus, we are firstly interested in minimizing the traveltime differences between data and synthetics. Future developments of this work may include, as well, minimization of an amplitude misfit function defined, e.g., in Tromp et al. (197) (2005).

Analogously to waveform tomography (section 2.3.2) and following Tromp et al. (197) (2005) and Tape et al. (187) (2007), we begin by defining a *traveltime misfit function* given, for example, by

$$F(\mathbf{m}) = \frac{1}{2}\sum_{r=1}^{N}\left[T_r^{obs} - T_r(\mathbf{m})\right]^2 , \tag{2.74}$$

where T_r^{obs} is the observed traveltime for the r-th source-receiver pair, and $T_r(\mathbf{m})$ the corresponding predicted traveltime for the model $\mathbf{m}$. Thus, the misfit function 2.74 is expressed in terms of the *cross-correlation traveltime difference*

$$\Delta T_r \ = T_r^{obs} - T_r(\mathbf{m}) , \tag{2.75}$$

that represents the measure of the misfit between data and synthetics chosen in this case of a traveltime tomography. A negative value of 2.75 corresponds to a delay in the synthetic arrival whit respect to the recorded arrival (Tape et al. (187), 2007). Another possible measurement in such a tomographic approach may be a *frequency-dependent multitaper traveltime difference* that leads to a *multitaper traveltime misfit function*, as shown in section 4.6.1. In a more general formulation, the misfit 2.74 may be weighted based upon the uncertainties on the considered measurements, as reported in section 2.3.4 and 4.6.1.

The variation of 2.74 due to a model perturbation $\delta\mathbf{m}$ is given by

$$\delta F = -\sum_{r=1}^{N} \Delta T_r \ \delta T_r , \tag{2.76}$$

where δT_r is the theoretical traveltime perturbation. As in a classical traveltime tomography based upon finite-frequency kernels (e.g., Luo & Schuster (121), 1991; Marquering et al. (131), 1999; Dahlen et al. (49), 2000), the variation δT_r may be related to the i-th component of the perturbed displacement, δs_i, due to $\delta \mathbf{m}$, by writing:

$$\delta T_r = \frac{1}{M_r} \int_0^T w_r(t) \partial_t s_i(\mathbf{x}_r, t) \delta s_i(\mathbf{x}_r, t) \, \mathrm{d}t \,, \tag{2.77}$$

where w_r denotes the cross-correlation window, and M_r the normalization factor

$$M_r = \int_0^T w_r(t) s_i(\mathbf{x}_r, t) \partial_t^2 s_i(\mathbf{x}_r, t) \, \mathrm{d}t \,. \tag{2.78}$$

One can substitute in 2.77 the Born approximation introduced for waveform tomography (equation 2.49), obtaining

$$\begin{aligned} \delta T_r = -\frac{1}{M_r} \int_0^T w_r(t) \partial_t s_i(\mathbf{x}_r, t) \int_0^t \int_V \big[\delta\rho(\mathbf{x}') \, G_{ij}(\mathbf{x}_r, \mathbf{x}'; t - t') \, \partial_{t'}^2 s_j(\mathbf{x}', t') \\ + \, \delta c_{jklm}(\mathbf{x}') \, \partial'_k G_{ij}(\mathbf{x}_r, \mathbf{x}'; t - t') \, \partial'_l s_m(\mathbf{x}', t') \big] \, \mathrm{d}^3\mathbf{x}' \mathrm{d}t' \mathrm{d}t \,. \end{aligned} \tag{2.79}$$

Then, using the reciprocity in 2.52, one can define the *traveltime (cross-correlation) adjoint source*

$$\bar{f}_i^\dagger(\mathbf{x}, t) = \frac{1}{M_r} \, w_r(T - t) \partial_t s_i(\mathbf{x}_r, T - t) \delta(\mathbf{x} - \mathbf{x}_r) \,, \tag{2.80}$$

which generates the *traveltime adjoint wavefield* [1]

$$\bar{s}_j^\dagger(\mathbf{x}', \mathbf{x}_r, T - t') = \frac{1}{M_r} \int_0^{T-t'} G_{ji}(\mathbf{x}', \mathbf{x}_r; T - t - t') \, w_r(T - t) \, \partial_t s_i(\mathbf{x}_r, T - t) \, \mathrm{d}t \,. \tag{2.81}$$

Compared to the waveform adjoint source 2.54, equation 2.80 refers to a single source-receiver pair r and it is independent of the measurements considered in this case, i.e., the traveltime differences. Moreover, the time-reversed signal that generates the adjoint filed at a given station is the i-th component of synthetic velocity at that station, $\partial_t s_i$. Given the adjoint field 2.81, the traveltime perturbation in 2.79, for an isotropic case, may be rewritten as

$$\delta T_r = \int_V \left(\bar{K_\rho} \, \delta \ln\rho + \bar{K_\mu} \, \delta \ln\mu + \bar{K_\kappa} \, \delta \ln\kappa \right) \mathrm{d}^3\mathbf{x} \,, \tag{2.82}$$

[1] The volume integral in 2.56 gives rise to the dependence of $\bar{s}_j^\dagger$ on $\mathbf{x}_r$ by using the properties of the delta function in 2.80.

with

$$\bar{K_\rho}(\mathbf{x},\mathbf{x}_r) = -\rho(\mathbf{x}) \int_0^T \bar{\mathbf{s}}^\dagger(\mathbf{x},\mathbf{x}_r,T-t)\cdot\partial_t^2\mathbf{s}(\mathbf{x},t)\,\mathrm{d}t\,, \tag{2.83}$$

$$\bar{K_\mu}(\mathbf{x},\mathbf{x}_r) = -2\mu(\mathbf{x}) \int_0^T \bar{\mathbf{D}}^\dagger(\mathbf{x},\mathbf{x}_r,T-t) : \mathbf{D}(\mathbf{x},t)\,\mathrm{d}t\,, \tag{2.84}$$

$$\bar{K_\kappa}(\mathbf{x},\mathbf{x}_r) = -\kappa(\mathbf{x}) \int_0^T \left[\nabla\cdot\bar{\mathbf{s}}^\dagger(\mathbf{x},\mathbf{x}_r,T-t)\right]\left[\nabla\cdot\mathbf{s}(\mathbf{x},t)\right]\,\mathrm{d}t\,, \tag{2.85}$$

where $\bar{\mathbf{D}}^\dagger$ is the traveltime adjoint strain deviator associated with $\bar{\mathbf{s}}^\dagger$. Alternatively, using (v_p, v_s, ρ) parametrization, one can rewrite (neglecting dependence on $\mathbf{x}$ and $\mathbf{x}_r$)

$$\delta T_r = \int_V \left(\bar{K'_\rho}\,\delta\ln\rho + \bar{K_\alpha}\,\delta\ln v_p + \bar{K_\beta}\,\delta\ln v_s\right)\,\mathrm{d}^3\mathbf{x}\,, \tag{2.86}$$

where

$$\bar{K'_\rho} = \bar{K_\kappa} + \bar{K_\mu} + \bar{K_\rho}\,, \qquad \bar{K_\alpha} = 2\left(\frac{\kappa+\frac{4}{3}\mu}{\kappa}\right)\bar{K_\kappa}\,, \qquad \bar{K_\beta} = 2\left(\bar{K_\mu} - \frac{4}{3}\frac{\mu}{\kappa}\,\bar{K_\kappa}\right)\,. \tag{2.87}$$

Equations 2.82 and 2.86 show that the traveltime variation can be recast from an integral over a time window w_r (eq. 2.79) to a volume integral involving 3D kernels. In particular, these equations can be rewritten as the following general relationship

$$\delta T_r = \int_V \bar{K_r}\,\delta\ln m\,\,\mathrm{d}^3\mathbf{x}\,, \tag{2.88}$$

that corresponds to the finite-frequency traveltime expression derived by Zhao et al. (210) (2000) and Dahlen et al. (49) (2000) in classical traveltime tomography. In 2.88 $\delta\ln m = \delta m/m$ denotes the generic relative perturbation of model parameters; $\bar{K_r}$ is the corresponding finite-frequency kernel, that, as highlighted by the subscript r, is specific for a given source-receiver pair, i.e. it reflects how a specific measurement in a seismogram, for a given source-receiver combination, "sees" the Earth model. Marquering et al. (131) (1999) called these kernels *banana-doughnut kernels*, due to their shape. The kernels derived so far (eqs. 2.83 - 2.85, 2.87) can be thought as the analogous, in adjoint tomography, of the "classical" banana-doughnut kernels, and they need to be constructed for each source-receiver pair via the interaction between the regular and the adjoint field.

As introduced in section 1.1, in realistic 3D problems, the calculations required to obtain all the banana-doughnut kernels increase dramatically and would become

almost unaffordable. However, one of the key points of the adjoint approach is that it is based upon the calculation of so-called *event and misfit kernels*, that, formally, may be constructed as sums of banana-doughnut kernels (eqs. 2.90 - 2.92 below), but, in practice, they are directly obtained from a single adjoint calculation for each event; thus, it is not necessary to calculate the banana-doughnut kernels for each source-receiver pair, as shown in what follows. Note that for waveform adjoint tomography in section 2.3.2, we have directly introduced the misfit kernels, whereas for traveltime adjoint tomography, we started by introducing the banana-doughnut kernels in order to draw connections with a classical finite-frequency traveltime tomography.

In order to introduce the mentioned misfit kernels, by substituting equation 2.86 in 2.76, one can obtain the variation of the traveltime misfit function in terms of 3D kernels (e.g., Tromp et al. (197), 2005)

$$\delta F = \int_V \left(K'_\rho \, \delta \ln \rho + K_\alpha \, \delta \ln v_p + K_\beta \, \delta \ln v_s \right) \, \mathrm{d}^3\mathbf{x} \,, \tag{2.89}$$

where, in this case, K'_ρ, K_α and K_β are the *traveltime misfit kernels* and are given by

$$K'_\rho(\mathbf{x}) = -\sum_{r=1}^{N} \Delta T_r \, \bar{K'_\rho}(\mathbf{x}, \mathbf{x}_r) \,, \tag{2.90}$$

$$K_\alpha = -\sum_{r=1}^{N} \Delta T_r \, \bar{K_\alpha}(\mathbf{x}, \mathbf{x}_r) \,, \tag{2.91}$$

$$K_\beta = -\sum_{r=1}^{N} \Delta T_r \, \bar{K_\beta}(\mathbf{x}, \mathbf{x}_r) \,. \tag{2.92}$$

Thus, the kernels 2.90 - 2.92, formally, are weighted sums of the banana-doughnut kernels in 2.87 over all the source-receiver combinations r, with weights represented by the corresponding traveltime difference ΔT_r for the r-th pair. Hence, these kernels include an explicit dependence on the data, i.e. on the traveltime measurements, whereas the banana-doughnut kernels are data independent. More generally one can write

$$\delta F = \int_V K \, \delta \ln m \, \mathrm{d}^3\mathbf{x} \,, \tag{2.93}$$

$$K(\mathbf{x}) = -\sum_{r=1}^{N} \Delta T_r \, \bar{K_r}(\mathbf{x}, \mathbf{x}_r) \,. \tag{2.94}$$

In this work (following e.g. Tape et al. (187), 2007), we use the term *event kernel* to denote the kernel referred to a single event, i.e., defined as the sum of the banana-doughnut kernels for an event over all the stations that recorded the event [1]. If the sum of the banana-doughnut kernels is considered over all the source-receiver combinations, as above, one obtains a *misfit kernel*, that, in fact, is simply the sum of the event kernels for all the considered events.

Based upon 2.81, one can now define the *combined traveltime adjoint field* (e.g., Tromp et al. (197), 2005)

$$\mathbf{s}^{\dagger}(\mathbf{x},t) = -\sum_{r=1}^{N} \Delta T_r \; \bar{s}^{\dagger}(\mathbf{x},\mathbf{x}_r,t)\,, \tag{2.95}$$

that is generated by the *combined traveltime adjoint source*

$$f_i^{\dagger}(\mathbf{x},t) = -\sum_{r=1}^{N} \Delta T_r \; \frac{1}{M_r}\, w_r(T-t)\, \partial_t s_i(\mathbf{x}_r, T-t)\, \delta(\mathbf{x}-\mathbf{x}_r)\,. \tag{2.96}$$

Thereby, to construct the adjoint field, each component of every receiver transmits simultaneously its signals in reverse time as fictitious sources (for traveltime tomography, the time-reversed predicted velocity in the time window w_r), each of which is located at the receiver and is weighted by the corresponding measurement ΔT_r. Thus, the key point is that now the adjoint source incorporates the measure of the misfit between data and synthetics, and, by using 2.96, the misfit kernels 2.90 - 2.92 can be directly calculated based upon only two simulations for each event, one involving the field $\mathbf{s}$ and the other one to propagate the adjoint field $\mathbf{s}^{\dagger}$ and to back-reconstruct $\mathbf{s}$. It is worth noting that, as we will show in section 2.3.4, the misfit kernels represent, in fact, the gradient of the misfit function necessary for the inversion, provided suitable basis functions have been introduced.

Figure 2.2, excerpt from Tape et al. (187) (2007), shows how a simple event kernel for a single source-receiver pair is constructed from the interaction between forward and adjoint fields, highlighting, then, the region of the model where data and synthetics differ.

To conclude this section, the key points of the adjoint approach are summarized in what follows, based upon the previous sections and, e.g., Tromp et al. (198) (2008).

[1] The expression of an event kernel is analogous to 2.94 provided the sum over r refers only to the stations that recorded the considered event.

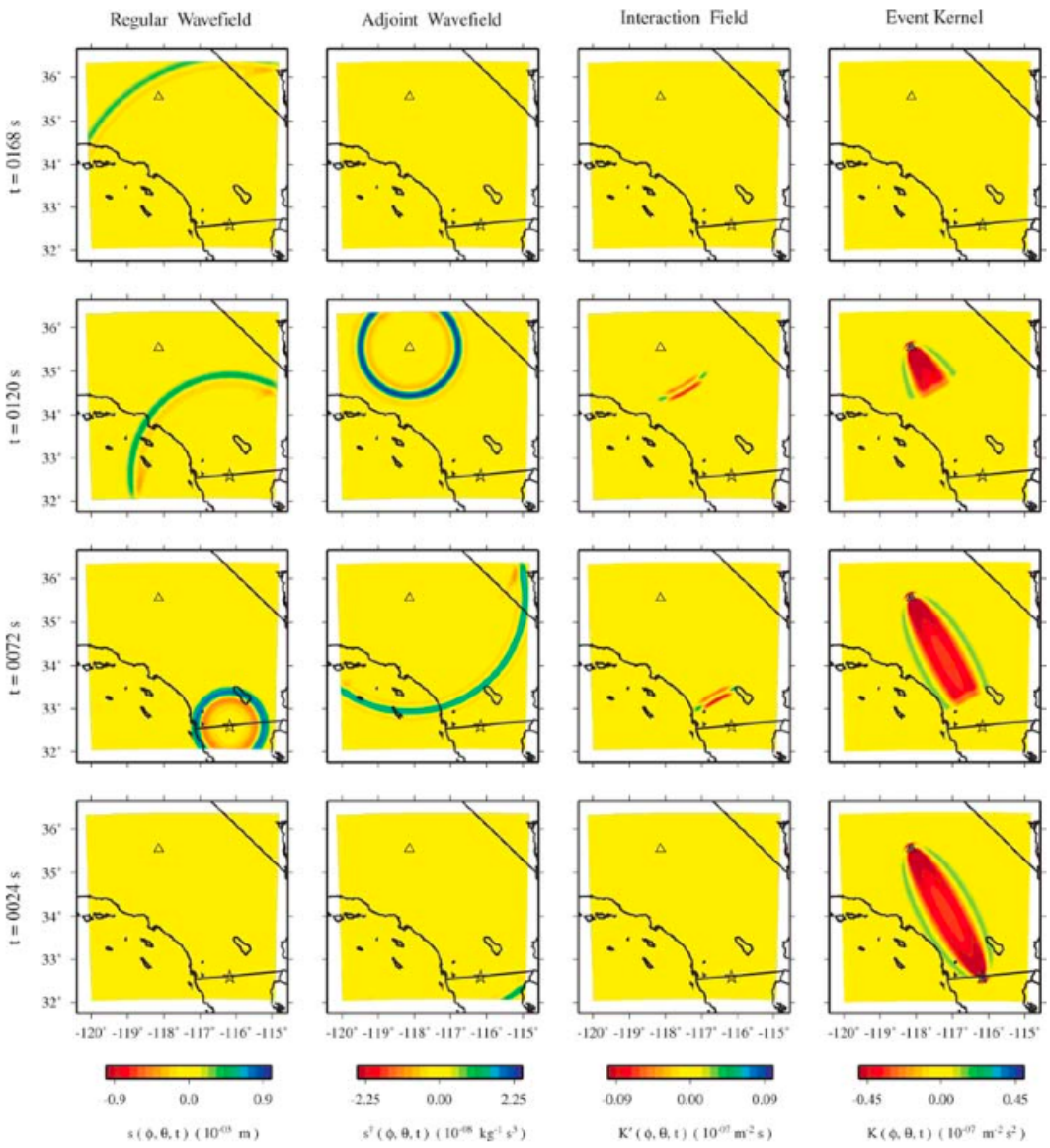


Figure 2.2: Traveltime event kernel - Sequence of interactions between the regular and adjoint wavefields during the construction of a traveltime cross-correlation event kernel. The star denotes the source, and the triangle denotes the receiver. Each row represents the time-step indicated on the left. In this case, with only a single receiver and a uniform model perturbation, the event kernel resembles a banana-doughnut kernel. The event kernel is constructed via the interaction between the forward wavefield (first column) and the adjoint wavefield (second column). The interaction field (third column) is the instantaneous product of the two wavefields, which is integrated to form the event kernel (fourth column). The event kernel shows the region of the current model that gives rise to the discrepancy between the data and the synthetics. Courtesy of Tape et al. (187) (2007).

- Adjoint tomographic inversions are based upon the calculation of 3D kernels, that are directly obtained from the interaction between the forward wavefield for the current model and the adjoint wavefield generated simultaneously at all the receivers. For each event, the construction of the kernels requires only one simulation for the regular field and one for the adjoint field. The overall misfit kernels, that highlight where the current model needs to be improved, are obtained by summing all the event kernels, resulting in the gradient of the misfit function required to solve the inverse problem (section 2.3.4). Thus, in general, at each step of the gradient-based minimization procedure, only $(3 * N_{events})$ 3D simulations (with N_{events} the number of events) are required (section 4.10), and the adjoint approach scales linearly only with the number of events, but it is independent of the number of the considered receivers, components and measurements. This is the main advantage with respect to a classical finite-frequency traveltime tomography, where, instead, a single banana-doughnut kernel for each measurement needs to be calculated. In an adjoint approach this would require a total of $(2N_{events}{}^*N_{receivers}{}^*N_{components}{}^*N_{picks})$ simulations, that is computationally unfeasible for 3D complex models. This issue is avoided by directly calculating event and misfit kernels. However, the calculation of single banana-doughnut kernels would allow one to construct, not only the gradient, but also the Hessian of the misfit function (section 2.3.4), as in a typical Newton method. Since this is generally not possible in an adjoint tomography for 3D models, gradient-based procedures need to be used in this tomographic approach and the initial model is improved by successive iterations.

- In the adjoint method the kernels are calculated on-the-fly by accessing simultaneously in memory to both the forward and the adjoint wavefields. This doubles the required computational memory, but avoids storing the Green's functions for all receivers and sources as a function of space and time (e.g., Zhao et al. (211), 2005). Details on how the kernel calculation is implemented in the spectral-element code used in this work are discussed in section 2.4.

- The kernels can be calculated for fully 3D heterogeneous models, that incorporate all the known complexities and that are considered, for example, in regional-scale seismology or in exploration seismology.

- In the seismograms recoded by each component of every receiver, for a given event, a large number of arrivals suitable as measurements can be present. In the adjoint approach, since it is independent on the number of picks, one is allowed to consider as much measurements as possible in the inversion, without the need to identify them. From the interaction between the forward and the adjoint fields one automatically obtains the 3D sensitivity associated to the pulses in the seismogram, that is represented by the kernels.

2.3.4 Iterative inversion strategy

To perform a tomographic inversion, as shown in general in section 2.3.1, one needs to calculate at least the gradient of the chosen misfit function. Thus, we begin by presenting the general formulations for the gradient and the Hessian in a traveltime finite-frequency tomography, to discuss, next, how the inverse problem can be solved in an adjoint tomographic approach.

Suppose to define a set of suitable basis functions B_k, k=1,...,M, in which any function of the problem can be expanded (e.g., Tape et al. (187), 2007). Thus, the relative model perturbation $\delta \ln m$ in 2.88 can be discretized based upon the B_k as

$$\delta \ln m(\mathbf{x}) = \sum_{k=1}^{M} \delta m_k \, B_k(\mathbf{x}) \,, \tag{2.97}$$

where δm_k, k=1,...,M, are the perturbed model coefficients in 2.45 - 2.46.

Substituting 2.97 in 2.88, one can rewrite the traveltime perturbation as

$$\delta T_r = \sum_{k=1}^{M} \delta m_k \, G_{rk} \,, \tag{2.98}$$

where

$$G_{rk} = \left. \frac{\partial T_r}{\partial m_k} \right|_{\mathbf{m}} = \int_V \bar{K}_r \, B_k \, \mathrm{d}^3\mathbf{x} \,, \tag{2.99}$$

are the elements of the so-called design matrix $\mathbf{G}$ and depend on the banana-doughnut kernels $\bar{K}_r$.

Analogously, by substituting 2.97 in 2.93, the variation of the misfit function becomes

$$\delta F = \sum_{k=1}^{M} \delta m_k \int_V K \, B_k \, \mathrm{d}^3\mathbf{x} \,. \tag{2.100}$$

Using the definition of the gradient of a misfit function, we then write

$$\delta F = \mathbf{g} \cdot \delta\, \mathbf{m} = \sum_{k=1}^{M} g_k\, \delta\, m_k\,. \tag{2.101}$$

Thus, comparing 2.100 and 2.101, we obtain the expression for the elements of the gradient $\mathbf{g}$

$$g_k(\mathbf{m}) = \left.\frac{\partial F}{\partial m_k}\right|_{\mathbf{m}} = \int_V K\, B_k\ \mathrm{d}^3\mathbf{x} \qquad k = 1, ..., M\,. \tag{2.102}$$

This equation is fundamental in view of the adjoint approach, since it shows that the gradient of the misfit function can be directly constructed based upon the misfit kernels K (as discussed later in more details). Following the classical approach, instead, the relationship between the misfit kernels and the banana-doughnut kernels 2.94 is substituted in 2.102 obtaining

$$g_k(\mathbf{m}) = -\sum_{r=1}^{N} \Delta T_r \int_V \bar{K}_r\, B_k\ \mathrm{d}^3\mathbf{x} = -\sum_{r=1}^{N} G_{rk}\ \Delta T_r \qquad k = 1, ..., M\,. \tag{2.103}$$

Defining the vector of the data (that in this case are cross-correlation traveltime measurements) as

$$\mathbf{d} = (\Delta T_1, ..., \Delta T_r, ..., \Delta T_N)^T\,, \tag{2.104}$$

the matrix form of the gradient in 2.103 is

$$\mathbf{g} = -\mathbf{G}^T\, \mathbf{d}\,, \tag{2.105}$$

where the dimension of the design matrix $\mathbf{G}$ is N x M. More generally, if the traveltime misfit function is weighted by the uncertainties associated to the traveltime measurements by introducing σ_r^2 in the denominator of 2.74, then, equation 2.105 becomes

$$\mathbf{g} = -\mathbf{G}^T\, \boldsymbol{C}_D^{-1}\, \mathbf{d}\,, \tag{2.106}$$

where the data covariance matrix $\boldsymbol{C}_D$ contains the variances σ_r^2 and it is diagonal. The values of the σ_r^2 will appear in the denominator of the adjoint sources, as well. Concerning the second derivative of the misfit function, that is the Hessian (eq. 2.43), its elements are

$$H_{kk'}(\mathbf{m}) = \left.\frac{\partial^2 F}{\partial m_{k'}\, \partial m_k}\right|_{\mathbf{m}} = \left.\frac{\partial g_k}{\partial m_{k'}}\right|_{\mathbf{m}} = \sum_{r=1}^{N} \left(G_{rk'}\, G_{rk} - \Delta T_r\ \left.\frac{\partial^2 T_r}{\partial m_k\, \partial m_{k'}}\right|_{\mathbf{m}} \right)\,. \tag{2.107}$$

The Hessian can be approximated by neglecting the second order terms and considering only the *approximate Hessian* given by

$$\tilde{H}_{kk'}(\mathbf{m}) = \sum_{r=1}^{N} G_{rk'}\, G_{rk} \qquad k, k' = 1, ..., M \ , \tag{2.108}$$

and in matrix form

$$\tilde{\mathbf{H}} = \mathbf{G}^T\, \mathbf{G}\,. \tag{2.109}$$

This approximation characterizes the so-called *Gauss-Newton methods.*

At this point, using the gradient 2.105 and the approximate hessian 2.109, the basic equation 2.45 of the inverse problem becomes, in a classical approach,

$$\mathbf{G}^T\, \mathbf{G}\, \delta\mathbf{m} = \mathbf{G}^T\, \mathbf{d}\,. \tag{2.110}$$

To determine the model update $\delta\mathbf{m}$ from 2.110 one needs to invert the Hessian (eq. 2.46), but the matrix 2.109 is in general not invertible. Thus, one can construct an approximate Hessian in the form (e.g., Tape et al. (187), 2007)

$$\tilde{\mathbf{H}}_\gamma = \mathbf{G}^T\, \mathbf{G} \ + \gamma^2\, \mathbf{D}\,, \tag{2.111}$$

where the damping matrix $\mathbf{D}$ (involving the norm, the gradient, or the second derivative of the model perturbations) and the damping parameter γ allow one to stabilize the inverse of the Hessian. Thus, finally, the inverse problem can be solved, in principle, by updating the model $\mathbf{m}$ with the perturbation $\delta\mathbf{m}$ calculated from

$$\delta\mathbf{m} = \left(\mathbf{G}^T\, \mathbf{G} \ + \gamma^2\, \mathbf{D}\right)^{-1} \mathbf{G}^T\, \mathbf{d}\,. \tag{2.112}$$

Our objective is to perform a tomographic inversion based upon an adjoint method, and, in such an approach, as already introduced in section 2.3.3, the calculation of the Hessian for 3D models is not feasible. Indeed, it would require the calculation of the banana-doughnut kernels for all the source-receiver pairs in order to construct the design matrix $\mathbf{G}$. Our inversion procedure is, thus, based just upon the gradient $\mathbf{g}$ of the misfit function. This gradient, in fact, can be readily computed from the 3D adjoint misfit kernels by using equation 2.102, instead of constructing $\mathbf{G}$ in equation 2.105 as in a classical tomographic approach. In section 2.3.3 we illustrated how the misfit kernels for traveltime adjoint tomography can be constructed. Now we need to define the basis functions that are required in the expression of the gradient 2.102. In

our work, the synthetic waveforms are simulated using the SEM (section 2.2), that discretizes the wavefields on a grid of Gauss-Lobatto-Legendre (GLL) points. Thus, a suitable choice would be to use the basis functions embedded in the numerical method (e.g., Tape et al. (187), 2007), i.e., the Lagrange polynomials defined, at the N_{GLL} GLL points, by equations 2.20 - 2.21. Each function $f(\mathbf{x})$ in the volume Ω can be expressed in terms of these basis functions as (see eq. 2.29)

$$f(\mathbf{x}) = \sum_{k=1}^{N_{GLL}} f_k \, L_k(\mathbf{x}), \tag{2.113}$$

where $f_k = f(\mathbf{x}_k)$ is the functional value at node $\mathbf{x}_k$, and $L_k(\mathbf{x})$ is a global function

$$L_k(\mathbf{x}) = \begin{cases} l_\alpha(\xi(x,y,z))\, l_\beta(\eta(x,y,z))\, l_\gamma(\zeta(x,y,z)) & \text{if } \mathbf{x}_k \in \Omega \text{ and } k|_\Omega = (\alpha,\beta,\gamma) \\ 0 & \text{if } \mathbf{x}_k \notin \Omega \end{cases} \tag{2.114}$$

with l_α, l_β, and l_γ Lagrange polynomials of degree α, β and γ, respectively. At the k-th node, considering 2.113, L_k must satisfy the relationship

$$L_k(\mathbf{x}_k) = 1\,. \tag{2.115}$$

The functions $L_k(\mathbf{x})$ are orthogonal but not orthonormal. Thus, we can construct a set of orthonormal basis functions $B_k(\mathbf{x})$ as

$$B_k(\mathbf{x}) = L_k(\mathbf{x})/\sqrt{V_k}, \tag{2.116}$$

where V_k is the volume associated with the k-th node:

$$V_k = \int_V L_k^2(\mathbf{x})\, \mathrm{d}^3\mathbf{x}. \tag{2.117}$$

Using $B_k(\mathbf{x})$, one can expand, for example, the misfit kernel $K(\mathbf{x})$

$$K(\mathbf{x}) = \sum_{k=1}^{N_{GLL}} \tilde{K_k} B_k(\mathbf{x}). \tag{2.118}$$

The expansion coefficients $\tilde{K_k}$ are determined by

$$\begin{aligned} \tilde{K_k} &= \int_V K(\mathbf{x}) B_k(\mathbf{x})\, \mathrm{d}^3\mathbf{x} \\ &= \int_V \sum_{k'} K_{k'} L_{k'}(\mathbf{x}) B_k(\mathbf{x})\, \mathrm{d}^3\mathbf{x} \\ &= \sum_{k'} K_{k'} \sqrt{V_{k'}} \int_V B_{k'}(\mathbf{x}) B_k(\mathbf{x})\, \mathrm{d}^3\mathbf{x} \\ &= K_k \sqrt{V_k} \end{aligned} \tag{2.119}$$

where $K_k = K(\mathbf{x}_k)$ is the value of the misfit kernel at a GLL grid point, and we have used 2.113, 2.116, and the orthogonality of functions B_k.

Now let us assume we have computed a misfit kernel $K(\mathbf{x})$. Since we are using SEM, it can be represented in a discrete form using its values $K(\mathbf{x}_k) = K_k$ at the N_{GLL} points $\mathbf{x}_k$. Thus, comparing equation 2.102 with 2.119, the discrete components of the gradient of the misfit function are given by

$$g_k = K_k \sqrt{V_k} \qquad k = 1, ..., N_{GLL}\,, \tag{2.120}$$

i.e., one obtains the key result that the value of the gradient at each GLL point $\mathbf{x}_k$ in the considered domain is just given by the misfit kernel evaluated at this point, provided the volume associated to the k-th node is defined.

2.3.4.1 Steepest descent minimization algorithm

In the formulations above we have derived the gradient of the misfit function for the adjoint approach, but, since the Hessian is not available in this case, equation 2.112 (or, in general, the Newton method 2.46) cannot be used to solve the inverse problem. Instead, the adjoint method has to be applied in combination with an iterative gradient-based minimization technique. We choose to use a *steepest descent algorithm* (e.g., Nolet (147), 1987), that is described in what follows.

Based upon the theory presented so far, given an initial model $\mathbf{m}^0$, one can calculate the misfit function $F(\mathbf{m}^0)$ (e.g., 2.74) from the differences between the data and the synthetics for model $\mathbf{m}^0$ (e.g., 2.75). Based upon these measurements, the sources of the adjoint wavefield can be constructed (e.g., 2.96), and, for each event, the interaction between regular and adjoint fields gives rise to the corresponding event kernel. Then, one calculates the misfit kernels, for the chosen model parametrization (e.g., 2.90 - 2.92), as a sum of the event kernels over all the considered events N_s:

$$K(\mathbf{x}) \;=\; \sum_{s=1}^{N_s} K^s(\mathbf{x})\,. \tag{2.121}$$

The initial gradient of the misfit function,

$$\mathbf{g}^0 = \mathbf{g}(\mathbf{m}^0) = \left.\frac{\partial F}{\partial \mathbf{m}}\right|_{\mathbf{m}^0}\,, \tag{2.122}$$

can be readily computed from the misfit kernels 2.121 by using 2.102 (or, practically, 2.120). It is worth noting that, after calculating the misfit kernels, some regularizations may be applied, such as a smoothing of the kernels (e.g., Tape et al. (187, 189), 2007, 2010) or a preconditioning (e.g., Tarantola (193), 2005). These aspects will be discussed in section 2.3.5.

At this point, in the steepest descent method, one sets the search direction to minimize the initial misfit function equal to minus the initial gradient of the misfit function. Thus, one defines the steepest descent vector [1]

$$\mathbf{p}^0 = -\mathbf{g}^0 , \tag{2.123}$$

and the first updated model is calculated as

$$\mathbf{m}^1 = \mathbf{m}^0 + \alpha \mathbf{p}^0 , \tag{2.124}$$

where α is a positive step length, and one implicitly assumes that the model parameters have been transformed into log-normalized quantities [2] (Tarantola (193), 2005). Then, the updated models for the following n iterations of the steepest descent method are given by

$$\mathbf{m}^n = \mathbf{m}^{n-1} + \alpha \mathbf{p}^{n-1} . \tag{2.125}$$

At each iteration, one needs to recalculate the corresponding misfit function and gradient (thus, the kernels) for the n-th model, and define $\mathbf{p}^n = -\mathbf{g}^n$. If $||\mathbf{p}^n|| < \epsilon$, with ϵ a suitable small quantity, then the model of the n-th iteration can be assumed as the solution of the inverse problem; otherwise, $\mathbf{m}^n$ is assumed as the new starting model and one continues iterating by using 2.125 in order to find the updated model that leads to the minimum of the misfit function.

The gradient of the misfit function gives us the minimization search direction (eq. 2.123). Then, as stated by equation 2.125, to calculate the model update $\delta\mathbf{m}^{n-1} = \alpha \mathbf{p}^{n-1}$, required to improve $\mathbf{m}^{n-1}$, one needs to determine the step length α, that

[1] Note that, if the uncertainties associated to the model parameters are introduced, this gives rise to a *model covariance matrix* $\boldsymbol{C}_M$ (Tarantola (193), 2005), that is diagonal if the model parameters are independent from each other. In this case, equation 2.123 becomes $\mathbf{p} = -\hat{\mathbf{g}}$, where the steepest ascent vector $\hat{\mathbf{g}}$ is calculated from the gradient by using $\boldsymbol{C}_M$, i.e., $\hat{\mathbf{g}} = \boldsymbol{C}_M \mathbf{g}$ (e.g., Tarantola (193), 2005).

[2] By labeling the i-th component of the log-normalized model vector $\mathbf{m}$ as m_i, the corresponding model parameter denoted by $\tilde{m}_i$ (e.g., v_p or v_s) has been transformed so that $m_i = \ln(\tilde{m}_i/\tilde{m}_{i_0})$, where $\tilde{m}_{i_0}$ is a scaling value for the considered parameter. See section 4.2.1.

quantifies how far one should go, along the search direction, to obtain a minimum of the misfit function. In order to do so, at each given iteration n, one can perform a line search by exploring the behaviour of the misfit function at this iteration for different values of α (e.g., Tarantola (193), 2005). In practice, one should select a subset of representative events and, fixing the initial model and the vector $\mathbf{p}$ to be those of the previous iteration, one estimates the new models $\mathbf{m}^n$ for a set of possible values of α by using 2.125 (see section 4.7.3). For each of these models, the corresponding misfit function is evaluated. One should find that the values of these misfit functions substantially describe a parabola as increasing α, thus, one should choose as steplength value the one that corresponds to the minimum of the misfit. This value is used in 2.125 to determine the new model $\mathbf{m}^n$ at iteration n.

2.3.4.2 Misfit formulations

To conclude this section, one should present the notation that we use when dealing with misfit functions (following J. Tromp, pers. com.).

As previously mentioned (section 2.3.3), in an adjoint approach, all the measurements made by comparing data and synthetics can be considered in the inversion. Thus, for any time window p in the time series, where data and synthetics have a significant amplitude and match reasonably well, one can estimate a measure of the difference between the two seismograms and, then, the corresponding misfit function F_p. For a given set of events N_s, data and synthetics can be compared, for example, in a chosen period band or for a chosen ground motion component (e.g., radial or transverse or vertical component). Thus, one says that they are compared for a specific *category* (e.g., SH waves on the transverse component) and one can, in general, consider any number of categories, for a total of N_c categories. The estimated value of the misfit function based upon all the measurements made for a given event s and a given category c may be denoted as F_{sc}, and it is given by

$$F_{sc} = \frac{1}{N_{sc}} \sum_{p=1}^{N_{sc}} F_p \,, \tag{2.126}$$

i.e., it is the sum of all the F_p for each window p, weighted by the total number of windows N_{sc} (suitable for a measurement) per event s and category c. Each function F_p could be weighted by a standard deviation σ_p and, if the data are fitted to within one

σ_p, $F_{sc} \approx 1$. Any type of misfit function can be considered. By using, e.g., the cross-correlation traveltime misfit function introduced in section 2.3.3 without the factor 1/2 and weighted by the cross-correlation uncertainties σ_p in each window, F_{sc} in 2.126 is given by

$$F_{sc} = \frac{1}{N_{sc}} \sum_{p=1}^{N_{sc}} \left(\frac{\Delta T_p}{\sigma_p} \right)^2 . \tag{2.127}$$

An analogous formulation can be obtained by using a multitaper traveltime misfit function (see section 4.6.1).

By summing over all the events, the misfit function F_c for a given category c is determined by

$$F_c = \sum_{s=1}^{N_s} w_s \, F_{sc} \, , \tag{2.128}$$

where w_s are the weights associated to each event such that $\sum_{s=1}^{N_s} w_s = 1$. In principle, every earthquake should contribute equally to the misfit, thus the weights are all equal to $1/N_s$. If the data are fitted to within one standard deviation, $F_c \approx 1$.

Finally, the overall misfit function F is obtained by summing over all the categories c, thus:

$$F = \sum_{c=1}^{N_c} w_c \, F_c \, , \tag{2.129}$$

where w_c are the weights associated to each category such that $\sum_{c=1}^{N_c} w_c = 1$. If every category has a similar contribution to the total misfit, the misfit functions F_c should be approximately equal to each other and the weights associated to each category are all equal to $1/N_c$. Again, if the data are fitted to within one standard deviation, $F \approx 1$.

2.3.5 Regularization: preconditioning and smoothing

Before using the kernels to represent the gradient needed in the inversion procedure, as mentioned in section 2.3.4.1, one can perform some regularizations on these kernels in order to stabilize and increase the convergence of the inversion procedure.

Preconditioning

One option is to apply a preconditioning to the sum of the event kernels, i.e., to the misfit kernels defined in 2.121. A possible choice for the preconditioner is a symmetric,

positive-defined function $P(\mathbf{x})$ given by (Y. Luo, pers. com.)

$$P(\mathbf{x}) = \int_0^T \partial_t^2 \mathbf{s}(\mathbf{x}, t) \cdot \partial_t^2 \mathbf{s}^\dagger(\mathbf{x}, T-t)\, \mathrm{d}t\,. \tag{2.130}$$

This preconditioner has an expression that is similar to one of the kernels defined above (cfr. eq. 2.67), except for the second temporal derivative of the adjoint field $\mathbf{s}^\dagger$. As a consequence, during the adjoint calculation, 2.130 can be computed on-the-fly, together with all the other kernels, from the interaction between the regular and the adjoint fields. For each event, thereby, we have a corresponding preconditioner P^s and, as the misfit kernels are calculated via 2.121, a combined preconditioner P can be obtained by summing the P^s over all the events N_s

$$P(\mathbf{x}) \;=\; \sum_{s=1}^{N_s} P^s(\mathbf{x})\,. \tag{2.131}$$

The preconditioned misfit kernels $\hat{K}$ are finally computed by multiplying the misfit kernels without preconditioning and the inverse of the the preconditioning matrix, i.e.

$$\hat{K}(\mathbf{x}) \;=\; P^{-1}(\mathbf{x})\, K(\mathbf{x})\,. \tag{2.132}$$

In order to avoid singularities in the inversion of the preconditioner, one usually assigns a threshold value to P when it is equal to zero. In general, the preconditioning is observed to reduce the effects of geometrical spreading on the shape of the kernels. The preconditioned kernels 2.132 can be used in 2.102 to calculate a new preconditioned gradient, that, used in the inversion procedure, should increase the convergence of the minimization algorithm (see section 4.7.3 and Fig. 4.7 (d)).

Smoothing

As proposed e.g. by Tape et al. (187) (2007), one can also apply a smoothing to the misfit kernels (with or without preconditioning). This is accomplished, for example, by convolving the kernels with a 3D Gaussian function given by

$$G(x,y,z) = \frac{1}{\left[(2\pi)^3\, \sigma_x^2\, \sigma_y^2\, \sigma_z^2\,\right]^{1/2}} \exp\left[-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2} + \frac{z^2}{2\sigma_z^2}\right)\right]. \tag{2.133}$$

Thus, the smoothed misfit kernel $\hat{K}$ is related to the unsmoothed one K by

$$\begin{aligned} \hat{K}(x,y,z) &= K(x,y,z) * G(x,y,z) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K(x',y',z')\, G(x-x', y-y', z-z')\, \mathrm{d}x'\, \mathrm{d}y'\, \mathrm{d}z'\,. \end{aligned} \tag{2.134}$$

In practice, the kernels are defined over a limited volume, thus 2.134 is rewritten as

$$\hat{K}(x,y,z) = \int_{z_1}^{z_2} \int_{y_1}^{y_2} \int_{x_1}^{x_2} K(x',y',z')\, G(x-x',y-y',z-z')\, \mathrm{d}x'\, \mathrm{d}y'\, \mathrm{d}z' \,, \quad (2.135)$$

and the convolution is performed at each GLL point in the mesh by using the smoothing function

$$G(x,y,z) = \frac{1}{N} \exp\left[-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2} + \frac{z^2}{2\sigma_z^2} \right) \right] . \quad (2.136)$$

The normalization factor N in 2.136 is given by

$$N \equiv \int_{z_1}^{z_2} \int_{y_1}^{y_2} \int_{x_1}^{x_2} \exp\left[-\left(\frac{x'^2}{2\sigma_x^2} + \frac{y'^2}{2\sigma_y^2} + \frac{z'^2}{2\sigma_z^2} \right) \right] \mathrm{d}x'\, \mathrm{d}y'\, \mathrm{d}z' \,, \quad (2.137)$$

such that

$$\int_{z_1}^{z_2} \int_{y_1}^{y_2} \int_{x_1}^{x_2} G(x,y,z)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z \; = 1 \,, \quad (2.138)$$

and N tends to the analytical expression $\left[(2\pi)^3\, \sigma_x^2\, \sigma_y^2\, \sigma_z^2\right]^{1/2}$ of equation 2.133 as the limits of the integrals tend to $\pm\infty$.

The smoothing operation is performed in order to reduce spurious amplitudes of the kernels in the immediate vicinity of sources and receivers, and to remove artificial effects due to the presence of short-wavelength unresolved features in the model (see section 4.7.2). The parameters σ in 2.133 and 2.136 are the half-width of the Gaussian (for the 3 directions) at a height of $e^{-1/2}G(0)$, thus they control the scalelength of the smoothing. In the adjoint inversion procedure, σ can be chosen based upon the shortest wavelength resolved in the simulation, i.e., by taking into account the size of the elements in the SEM mesh. Thus, the smoothing gives an indication of the resolution of our model and one can reduce its scalelength as the model improves at successive iterations.

An alternative approach, not used in this work, is to smooth the model by adding a damping term to the misfit function (e.g., Akçelik et al. (1, 2), 2002, 2003).

2.4 SPECFEM3D Version 2.0 'Sesame'

The high-performance computing power is dramatically and quickly increasing. Thus, it is of crucial importance to develop softwares that are able to exploit these advantages in order to improve capabilities of simulating seismic wavefields and imaging Earth's structure and rupture processes. In this work, we have used the new version, 2.0, of

the code SPECFEM3D, also called 'Sesame' [1], that is presented in details in the user manual [2] and in Peter et al. (160) (2011). In this section, I will summarize the structure and the main features of the code that make it to be very well suited to address both forward and inverse problems in seismology. Moreover, I will highlight the specific implementation settings used in this thesis.

The code is based upon the spectral-element method presented in section 2.2; thus, it takes advantage of the powerful characteristics of such a technique, in order to accurately simulate seismic wave propagation in highly complex, 3D heterogeneous media. In particular, it can be successfully applied to local and regional scale problems, featuring strong velocity contrasts and/or distorted geometries, as shown in Peter et al. (160) (2011). Moreover, thanks to the implementation of the adjoint method (section 2.3), in combination with SEM, the software has an essential role in seismic inversion procedures, since it allows one to simulate adjoint wavefields and to calculate the required kernels by exploiting the powerful capabilities of the adjoint approach. SPECFEM3D 2.0 is able to work both in (an)elastic and acoustic domains, implementing the coupling conditions between the two, as well. The study of the Earth 'noise' can be also addressed to (Peter et al. (160), 2011), by enabling one to calculate noise sensitivity kernels (Tromp et al. (200), 2010). However, this tool of the code has not been used in this work.

Concerning forward simulations of seismic waveforms in the (an)elastic case (in which we are interested), the weak formulation of the equation of motion 2.11, i.e. equation 2.17, is solved based on the spectral-element theory presented in section 2.2.

The first crucial step is, thereby, the construction of a high quality mesh for the considered region. One option is to use the internal mesher in SPECFEM3D, called meshfem3D. It allows to design relatively simple meshes for layercake models, using an analytical linear interpolation from the top to the bottom of the mesh (Peter et al. (160), 2011). For more complex geometries with lateral heterogeneities, as in our case, one can alternatively use the advanced 3D unstructured hexahedral mesh generator CUBIT (Blacker et al. (17), 1994). Meshes constructed with this external tool can discretize volumes of arbitrarily complex shapes and they can honor all internal discontinuities in the model as well as the free surface topography. Moreover, a 3D tomographic model

[1] freely available at http://www.geodynamics.org/svn/cig/seismo/3D/SPECFEM3D

[2] see http://www.geodynamics.org/svn/cig/seismo/3D/SPECFEM3D/trunk/doc/USER_MANUAL

of the considered region can be accounted for by using mesh elements of suitable size. The construction of high quality 3D meshes need to be performed by experts and may require a very long time. In this work, we used GEOCUBIT[1], a Python script collection based upon CUBIT (Casarotti et al. (24), 2008), to built 3D hexahedral meshes for central Italy (sections 3.4.2 and 4.2.2), by satisfying the mesh design requirements in section 2.2.2 and using hexahedral elements with 8 control points. Examples of fully unstructured hexahedral meshes constructed based on CUBIT are shown in Figure 2.3.

When using meshfem3D, the partition of the mesh into slices, to be assigned to different processors for the parallel implementation (section 2.2.4), is automatically performed by the mesher. On the contrary, if the mesh is built by using CUBIT, then a suitable partioner is required. An example of external partitioner, used in this work, is the software package SCOTCH (Pellegrini & Roman (158), 1996; Chevalier & Pellegrini (37), 2008), that allows a load-balanced partitioning of complex meshes on an arbitrary number of cores. An alternative partitioner able to accomplish these tasks is METIS (Karypis & Kumar (91), 1998), but SCOTCH is more actively maintained (Chevalier & Pellegrini (37), 2008) and performs better in many cases (Peter et al. (160), 2011). An example of mesh partitioning and load-balancing by using SCOTCH is shown in Figure 2.4.

After meshing and partitioning, the step required before performing the simulations with SPECFEM3D is to generate mesh databases for each partition (Peter et al. (160), 2011). These databases contain the GLL points (section 2.2.3) for all the spectral elements and they represent an input for the SEM solver. Based upon the considered models, material properties are assigned to these GLL nodes. In particular, we contributed to develop the code so that any external 3D tomographic model can be read in and, based upon an interpolation algorithm, the values of model parameters are attributed to each GLL points.

The last input to run the simulations are the definition of the location of receivers and the representation of seismic sources. In particular, one can implement a point source represented by a Gaussian source time function, or a source of finite size described as a superposition of Gaussian point sources (e.g., sections 4.3 or 3.4.3, respectively). The parameters required by the code for each point source are the location, the

[1]freely available at http://www.geodynamics.org/svn/cig/seismo/3D/GEOCUBIT

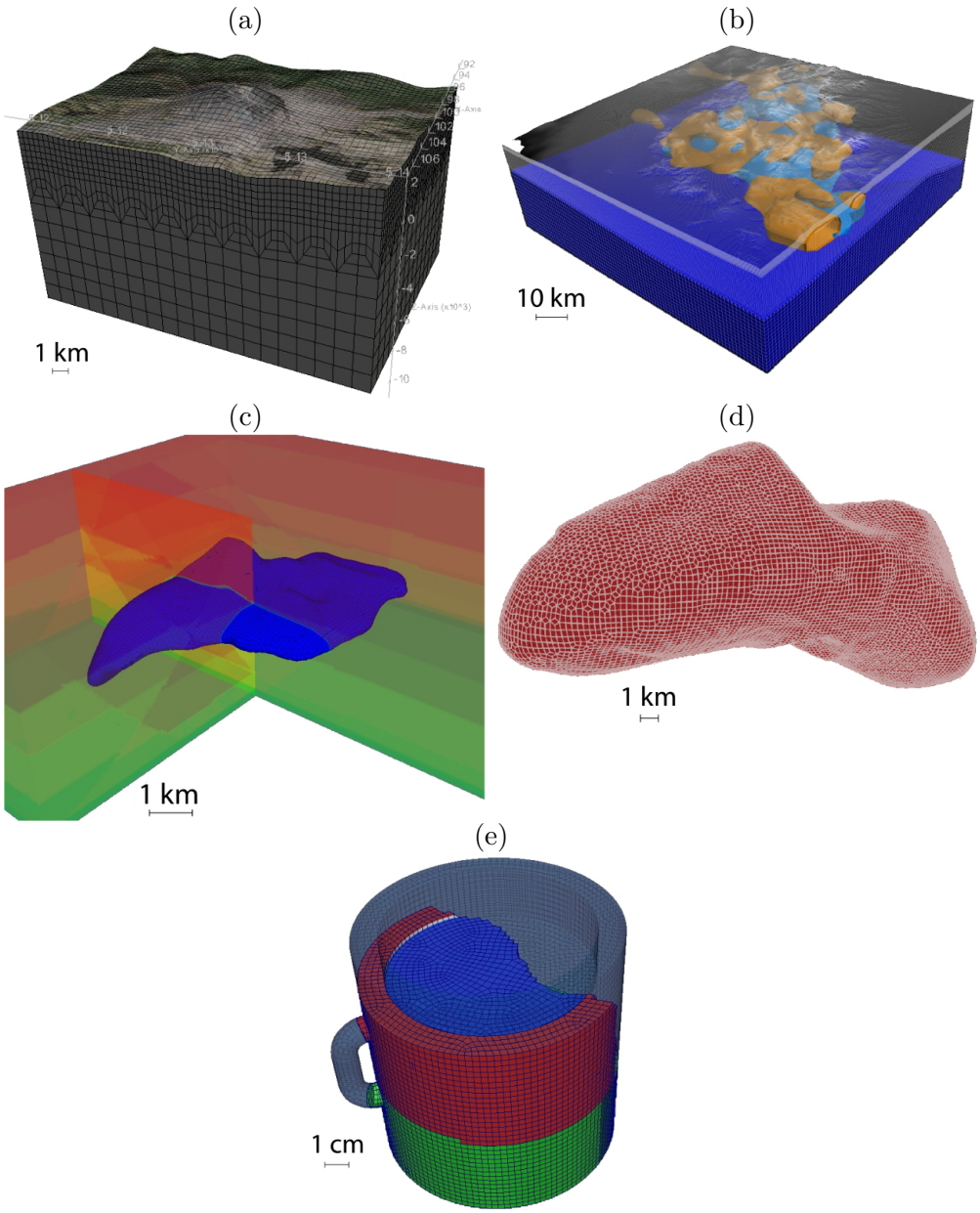


Figure 2.3: Mesh examples - (a) Mount St. Helens meshed by hexahedral elements. The mesh honors surface topography and includes a mesh tripling layer in the middle of the model. The smallest element size is approximately 280 m. (b) L'Aquila, Italy, region discretized for high-frequency simulations. The mesh honors surface and Moho topography and includes two mesh tripling layers. The orange and blue volumes denote slower and faster than average wavespeeds, respectively (see section 3.4.1 for details). (c) Salt dome body meshed inside an exploration model for a SEG/EAGE benchmark test. (d) 3D hexahedral mesh of the asteroid 433-Eros. (e) Arbitrarily-shaped mesh for coupled solid-fluid simulations involving a coffee cup. Courtesy of Peter et al. (160) (2011).

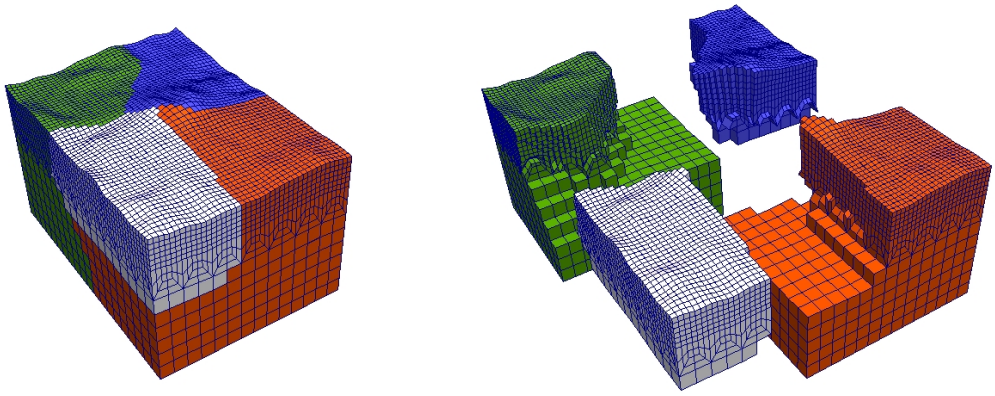


Figure 2.4: Mesh partitioning - Mount St. Helens mesh partitioned and load balanced (by using SCOTCH) to run in parallel on four cores. The four partitions are indicated by different colors. Courtesy of Peter et al. (160) (2011).

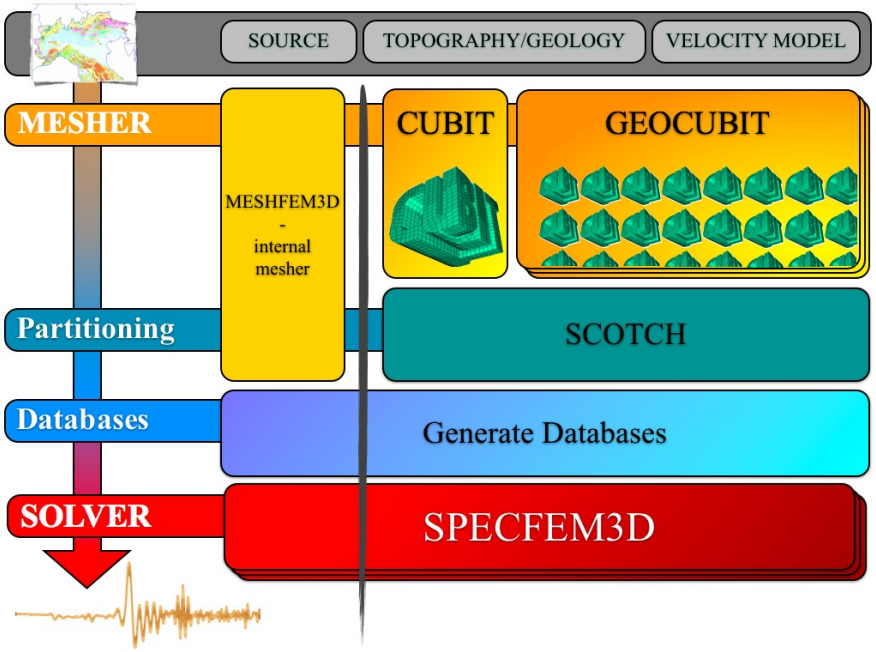


Figure 2.5: SPECFEM3D 2.0 Workflow - Workflow for running spectral-element simulations with SPECFEM3D Version 2.0 'Sesame'. The gray box on top contains the input elements for the code.

6 components of the moment tensor and the half duration and origin time of the rupture. At this point, the spectral-element solver performs a numerical integration of the wave equation, simulating the synthetic waveforms for each of the considered stations. Fixed the region and the resolution, one does not need to remesh the domain for every simulation. The main steps in using SPECFEM3D 2.0 are summarized in Figure 2.5.

It is worth noting that the code can also account for the attenuation effects on wavefields (Savage et al. (172), 2010) based upon the formulation in section 2.1.1 and using (in our case) three standard linear solids. It can accommodate full 21-parameter anisotropy, as well (Chen & Tromp (35), 2007). Moreover, SPECFEM3D implements absorbing conditions at fictitious boundaries of the mesh based upon the formulation of Clayton & Engquist (44) (1977) and Quarteroni et al. (167) (1998) (see eq. 2.16). On the contrary, a well load-balanced parallel implementation of the code when using Perfectly Matched Layer conditions (section 2.1.2) is still challenging (Peter et al. (160), 2011). Finally, in our work we set the code so that the degree of the polynomials used by SEM to discretize wavefields on elements (section 2.2.3) is equal to 4, thus the edges of each hexahedral element contain 5 unevenly spaced GLL points (see Fig. 2.1).

As mentioned at the beginning of this paragraph, one can also use SPECFEM3D to construct the adjoint kernels that are needed to solve the inverse problem. In section 2.3 we showed that these kernels are obtained from the interaction between the regular displacement field $\mathbf{s}$ for the current model and the adjoint field $\mathbf{s}^\dagger$, that is generated simultaneously at all the receivers based upon time-reversed signals. As a consequence, one needs to access at the same time to both the forward and the adjoint fields, at instant t and $T-t$ respectively. To this end, using the code SPECFEM3D, it is possible to run, for any event, a first forward simulation and to save the displacement at the last time step of this simulation, i.e., the final displacement $\mathbf{s}(\mathbf{x}, T)$ [1]. Then, the adjoint simulation for the given event is launched and it involves both the advance of the adjoint field $\mathbf{s}^\dagger$, and the back-reconstruction of the regular field by using $\mathbf{s}(\mathbf{x}, T)$ as a starting point for an integration of $\mathbf{s}$ backward in time (e.g., Tromp et al. (197, 198), 2005, 2008). This implies that an adjoint simulation requires almost twice the computational time of one forward simulation, since it involves both $\mathbf{s}$ and $\mathbf{s}^\dagger$.

[1] In general, all the state variables for the last time step are saved, including wavefields of displacement, velocity and acceleration.

Based upon the procedure above, using SPECFEM3D, one runs, for each event, one forward and one adjoint simulation, where the input sources are the real event and the combined source in 2.96, respectively. Hence, from the interaction of $\mathbf{s}^\dagger$ with the reconstructed $\mathbf{s}$, both simultaneously carried in memory, the code calculates on-the-fly the event kernels 2.90 - 2.92 for each event. We contributed to implement additional routines in the code to obtain the misfit kernels, by summing the event kernels, to precondition and smooth them, and to perform the model update based upon eq. 2.125.

In general, the code can calculate the kernels for any given pulse in a seismogram, i.e., it can be also used to calculate the finite-frequency banana-doughnut kernels 2.87 for every selected time window, provided the code receives as input of the adjoint simulation an adjoint source such as 2.80. Figure 2.6 (excerpt from Peter et al. (160), 2011) shows an example of the elastic, isotropic kernels 2.83-2.85 and 2.87, obtained by using SPECFEM3D 2.0. Figure 2.7 shows the misfit kernels 2.91 and 2.92 obtained for central Italy region by summing the event kernels for several events including foreshocks and aftershocks of the 2009 L'Aquila earthquake (see chapter 4).

Note that, as reported in Tromp et al. (197) (2005), the backward integration of the regular field $\mathbf{s}$ also involves an undoing of the attenuation effects, that is considered a numerically stable process (Tarantola (192), 1988). Moreover, for models with absorbing boundaries, one needs to store $\mathbf{s}(\mathbf{x}, t)$ on these boundaries as a function of time in order to back-reconstruct $\mathbf{s}$ (Gauthier et al. (78), 1986).

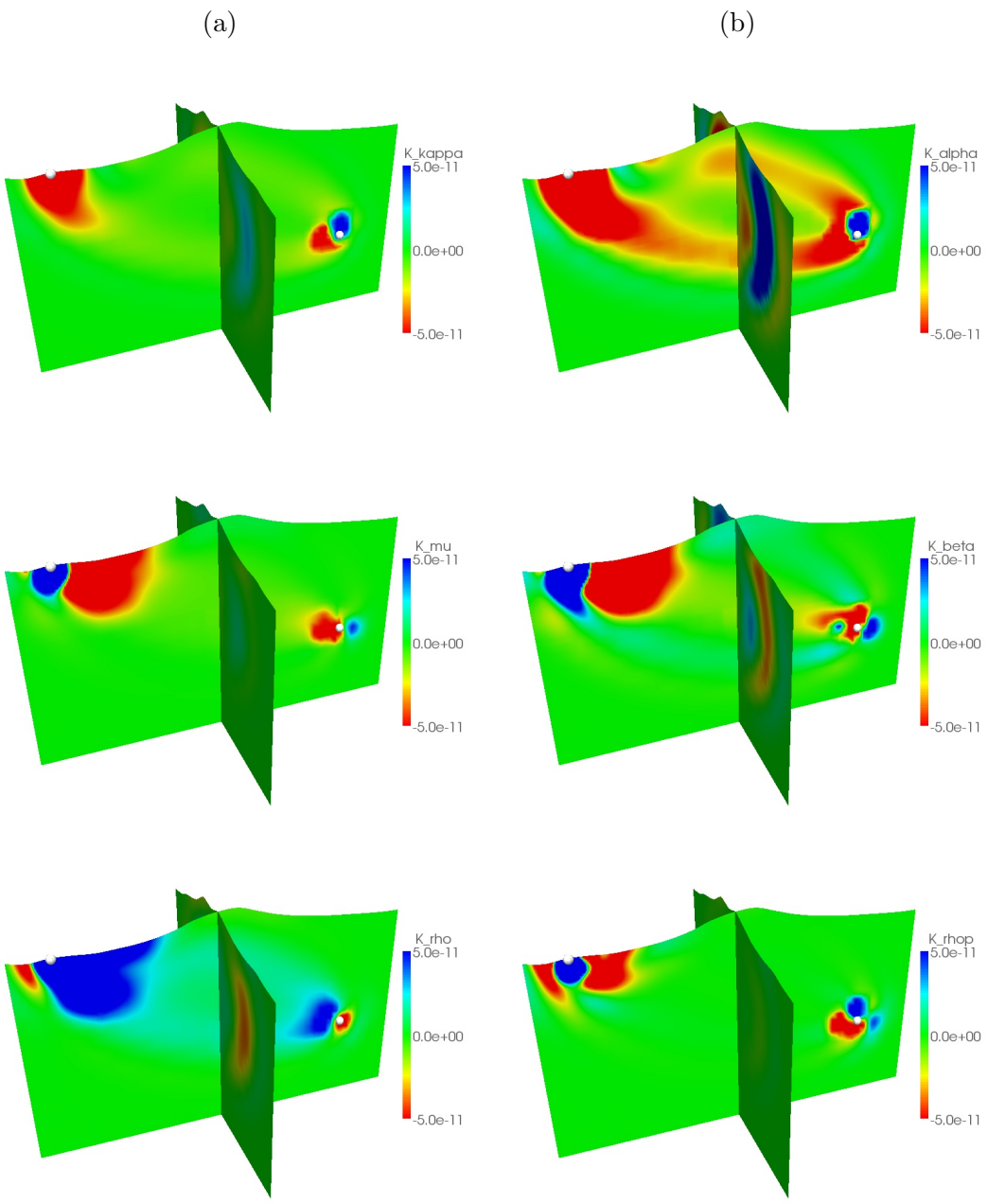


Figure 2.6: Traveltime phase-specific sensitivity kernels - Fréchet derivatives for isotropic parameterizations (a) K_κ, K_μ & K_ρ and (b) K_α, K_β & K'_ρ are compared in a model of Mount St. Helens using traveltime adjoint sources for the P wave. Shown are vertical cross sections through the source-receiver (white dots) line and perpendicular to this line. Courtesy of Peter et al. (160) (2011).

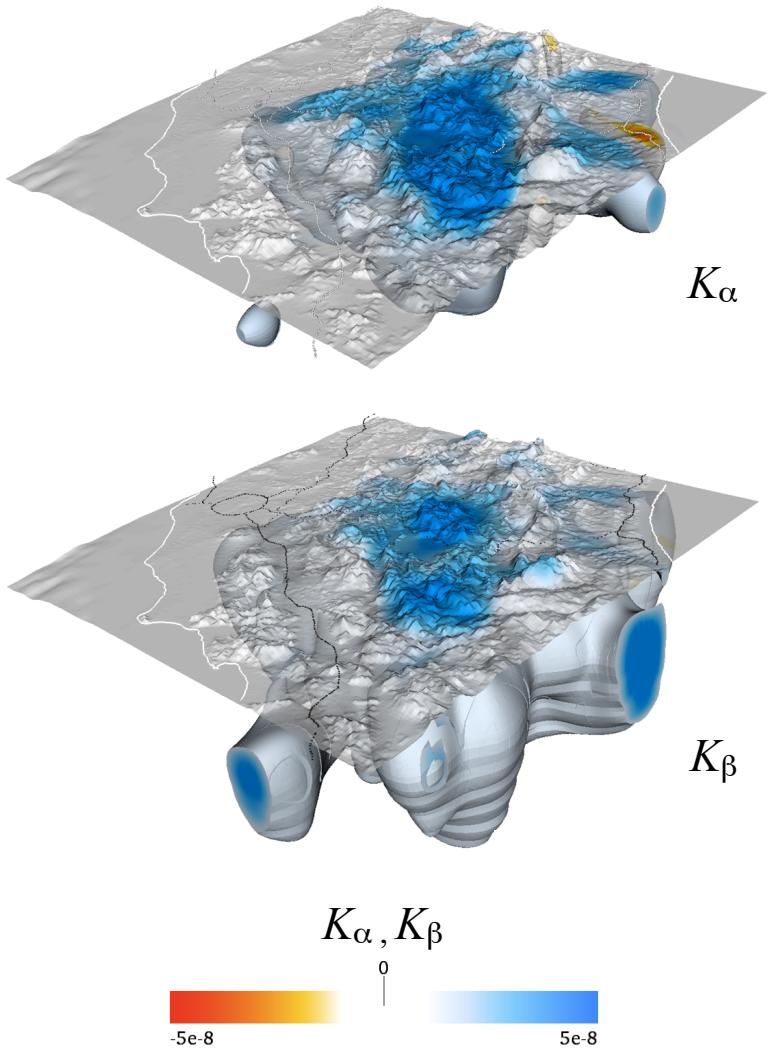


Figure 2.7: Traveltime misfit kernels - Fréchet derivatives K_α and K_β of the traveltime misfit function used in the tomographic inversion for central Italy (sections 4.7.1 - 4.7.2). They are obtained by summing the corresponding event kernels over all the considered events and applying some regularizations. The scale is in m^{-3}.

Chapter 3

SEM simulations of seismic waves generated by the 2009 L'Aquila earthquake

3.1 Introduction

The spectral-element method (SEM) represents a very accurate and performant numerical technique (section 2.2). It allows us to simulate seismic wavefields for complex realistic structures, accounting for all aspects that influence 3D wave propagation, included a kinematic description of sources of finite size. Numerous examples of successfully SEM applications are present, both for global, regional and local scale problems (e.g., Komatitsch et al. (110), 2005; Tape et al. (189), 2010; see section 1.1 for more references). As a consequence, we decided to apply this method to study seismic waveforms, that have been generated by the 2009 M_w 6.3 L'Aquila (Italy) mainshock and that propagated in central Italy. This chapter will discuss the results obtained by considering a finite fault model, and both 1D velocity profiles (with and without topography) and a more realistic 3D tomographic model for the considered region. The theoretical formulation of the problem is explained in section 2.2. The SE code used for the simulations, SPECFEM3D 2.0, has been presented in section 2.4.

The 2009 L'Aquila event offers a great opportunity, by applying SEM in this local context, to investigate the response of a highly complex realistic medium to a relatively large earthquake. Our study is one of the first examples that takes into account, in

the simulations, the complexities of both the kinematic rupture and the heterogenous structure, and compares the resulting synthetics and the data.

The considered mainshock occurred in a densely populated area of the central Apennines and causes more than 300 fatalities (e.g., Scognamiglio et al. (174), 2010). An unprecedented amount of high-quality data have been recorded (e.g., Akinci et al. (5), 2010), including seismograms from permanent velocimetric stations and accelerometric stations (section 3.3).

Moreover, central Italy is characterized by a very complex heterogeneous structure (section 3.2). Many low velocity areas, mostly in correspondence of sedimentary basins, feature shallow surface layers, together with remarkable topographic reliefs, that belong to the central sector of the Apennines. Also the depth velocity profile is highly complex, with remarkable lateral variations, and the Moho discontinuity features a complicated topography, generally deeper toward the Adriatic side. Numerous studies are available, inferring these structure complexities (e.g., Ponziani et al. (163), 1995; Alessandrini et al. (6), 1995; Chiarabba & Amato (38), 1996; Di Stefano et al. (58, 59, 60), 1999, 2009, 2011; Mele & Sandvol (139), 2003; Li et al. (116), 2007; Chiarabba et al. (39), 2010). Moreover, the SE code allows us to describe this region with meshes of hexahedral elements that incorporate all the known heterogeneities and any type of model discontinuity. We note that these 3D meshes could enable us to simulate seismic wavefields in the region at very high frequencies (up to $\sim$5 Hz, section 3.4.2).

Several recent studies present results of source inversions performed to infer source mechanism and rupture process of the M_w 6.3 L'Aquila earthquake (e.g., Cirella et al. (42), 2009; Atzori et al. (10), 2009; Scognamiglio et al. (174), 2010; Cheloni et al. (34), 2010). Thus, we have at our disposal many valuable information to constrain the source model. In particular, for this work we use the finite source description of Cirella et al. (42) (2009), mainly because it was the only one (when we started our work) obtained by a non-linear joint inversion of both strong motion and GPS data. SPECFEM3D enables us to study the finite-frequency effects on the wavefield due to a fault of finite size, that is implemented in the code as a combination of point sources (section 3.4.3).

Finally, implications for seismic hazard assessment are possible from the study of L'Aquila event. In fact, using SPECFEM3D, one can also construct synthetic peak ground velocity (PGV) maps, that provide indications of the strong ground shaking caused by the mainshock (section 3.8).

The next sections are organized as follows.

- General description of the simulated region and of the 2009 L'Aquila mainshock (section 3.2).
- Presentation of the analysed seismic data (section 3.3) and of the velocity and source models used for the SE simulations with SPECFEM3D (section 3.4).
- Discussion of the results obtained from simulations of L'Aquila seismic wavefield using 1D models with and without the surface topography (section 3.6). This will highlight the SEM capability of implementing topographic complexities on synthetic waveforms. Moreover, comparing data and synthetics, it will show ability and limits of our 1D models in capturing features in the seismograms related to topography.
- Discussion of the results obtained by using a 3D velocity model, with respect to a 1D profile (section 3.7). This shows once again that SEM is perfectly suitable to implement all the complexities that affect realistic wave propagation. Moreover, the importance of considering a 3D structure, together with the limits of the available velocity and source models, will be presented. An analysis within different frequency bands, will give an indication of the problem accuracy.
- Presentation of peak ground velocity synthetic estimates based upon both 1D and 3D velocity models of the region (section 3.8).

3.2 L'Aquila earthquake and the simulated region

On April 6th 2009, at 01:32 UTC, an earthquake with moment magnitude M_w=6.3 occurred in the Abruzzi region, ∼5 km SW from the city of L'Aquila. It ruptured a normal fault striking NW-SE along the central Apennines axis and dipping at about 50° to the SW. The hypocenter location, revised by INGV[1],is 42.35°N, 13.38°E, with a depth of 9.5 km. Despite its moderate size, the event caused about 300 casualties and severe damages in the city of L'Aquila and the surrounding villages.

[1]http://portale.ingv.it/primo-piano/archivio-primo-piano/notizie-2009/terremoto-6-aprile/copy_of_la-sequenza-sismica-dell-aquilano-aprile-2009/

A sequence of foreshocks started few months before the main event, with the largest earthquake that occurred on March 30th with M_w=4 (e.g., Scognamiglio et al. (174), 2010). Several aftershocks followed the April 6th event, among them the M_w=5.6 event of April 7th and the M_w=5.4 event of April 9th are the largest (Fig. 3.1 (b)).

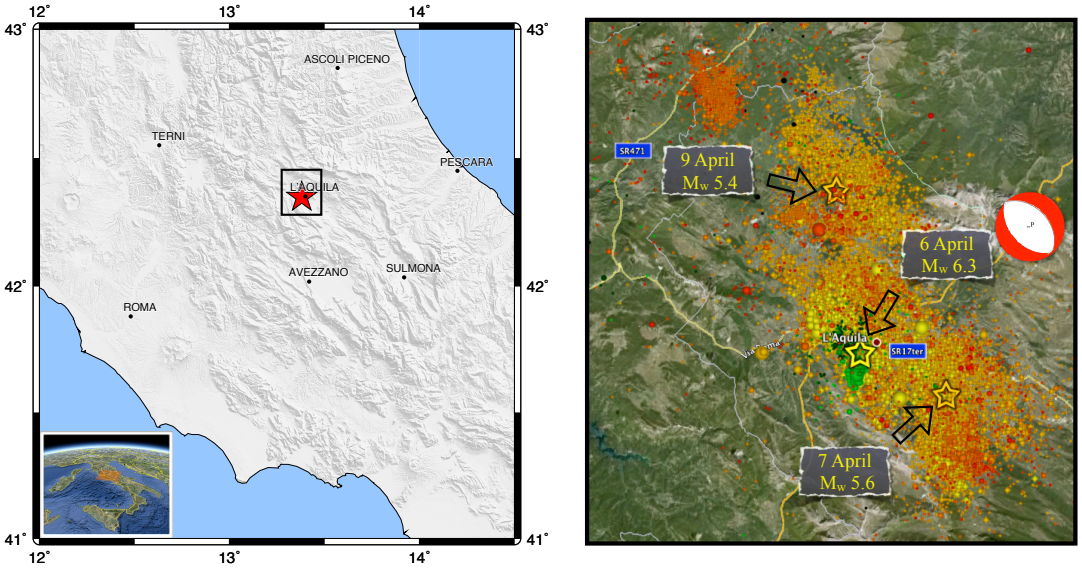


Figure 3.1: Simulated region and epicentral area of the 2009 L'Aquila event - (a) Area in central Italy for the study of the wavefield generated by the 2009 L'Aquila event (red star). (b) Close-up of the black inset in fig.(a) that includes the epicentral area. Location and focal mechanism of the main event are indicated, together with its foreshocks (green) and aftershocks in this area.

In this work we are interested in studying the seismic wavefield generated in central Italy by the L'Aquila mainshock. The considered region features a complex tectonic structure and evolution (e.g., Malinverno & Ryan (128), 1986; Patacca & Scandone (154), 1989; Doglioni (61, 62), 1991, 1995; Scrocca (175), 2006), reflected by the high variability of geologic and geomorphologic patterns. At present, the central sector of the Apennines chain is characterized by a NE-trending horizontal extension (e.g., Mariucci et al. (130), 1999; Hunstad et al. (85), 2003; Montone et al. (143), 2004; Li et al. (116), 2007), related to the opening of the Tyrrhenian back-ark basin (e.g., Malinverno & Ryan (128), 1986). This extensive regime causes the formation of intramountain basins, such as L'Aquila, Sulmona and Fucino basins, covered by Plio-Quaternary continental

sediments (e.g., Cavinato & De Celles (26), 1999; Chiarabba et al. (39), 2010, Fig. 2). Thus, most of the active faults are normal, composing a fault system striking NW-SE along the Apennines and bounding the basins (e.g., Chiaraluce et al. (40), 2004; Patacca et al. (155), 2008). In particular, the L'Aquila event epicentral area corresponds to the upper and middle Aterno valley, with Quaternary lacustrine deposits forming its basins. The depth of the deposits in this area varies from about 60 m to more than 200 m from the upper to the middle Aterno valley (Bosi & Bertini (18), 1970), and many evidences of ground motion amplifications have been found (e.g., De Luca et al. (54), 2005; Akinci et al. (5), 2010). Finally, remarkable topographic ridges characterize the considered region, such as the Gran Sasso and Maiella massifs to the West and the Simbruini Mts to the East (e.g., Chiarabba et al. (39), 2010, Fig. 2).

The volume involved in our simulations extends 200 km x 200 km horizontally and 60 km in depth in central Italy, with limits of latitude and longitude, respectively, (41.10-42.90)°N and (12.04-14.45)°E. Figure 3.1 (a) shows the target region, and the close-up in subfigure (b) represents the area around the L'Aquila mainshock, together with many of the foreshocks and aftershocks.

To describe the main characteristics of the L'Aquila event we refer to the study of Cirella et al. (42) (2009). They adopted a fault plane with a strike of N133°E and a dip of 54° to the SW. The strike direction is taken from the results of SAR interferometric data analysis (Atzori et al. (10), 2009). The dip value is consistent with both the hypocenter location and the observed surface breakages (EMERGEO (68) Working Group 2010). In addition, the assumed strike and dip are within the range of values inferred from moment tensor solutions[1]. Aftershock distribution and results on GPS displacements (Anzidei et al. (9), 2009) also contributed to identify the fault geometry.

We rely on the study of Cirella et al. (42) (2009) mainly because their source inversion involves also seismic strong motion waveforms (together with GPS data). Moreover, using the inverted model, they also provide synthetic timeseries that can be compared with our simulations. The principal results of Cirella et al. (42) (2009)'s work show a heterogeneous distribution of the slip on the fault plane. Two main patches are evident, a shallow small slip feature, located up-dip from the hypocenter, and a larger deeper (between 9 and 14 km depth) feature, located southeastward. This slip distribution is in agreement with the on-fault aftershock pattern and the induced

[1] http://www.emsc-csem.org

surface breakages. Moreover, they infer a larger rupture velocity in the up-dip direction compared to the along-strike direction. This is probably related to a lower velocity layer imaged by the crustal profile used by Cirella et al. (42) (2009) at depths corresponding to the largest slip patch. Figure 3.2 (b) shows the fault attributed to the L'Aquila mainshock, projected on the surface, and the corresponding slip distribution inferred from the source inversion of Cirella et al. (42) (2009).

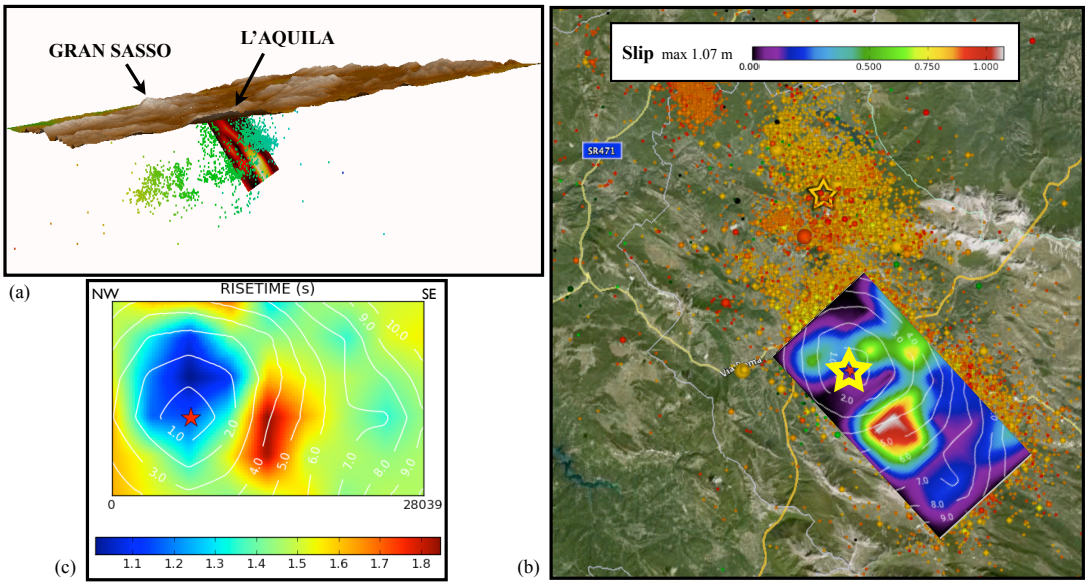


Figure 3.2: Source model of L'Aquila event - (a) Fault plane viewed from NW. The red sphere on the plane is the nucleation point, the green squares are the aftershocks of the following 10 days. Superimposed is the topography (section 3.4.1). (b) Same region in central Italy as Fig. 3.1 (b) with the fault plane of Cirella et al. (42) (2009) projected on the surface. The red star indicates the mainshock hypocenter. The colors describe the on-fault slip distribution inferred by Cirella et al. (42) (2009), the rupture time is shown by white contour lines (in sec). (c) On-fault rise time distribution (Cirella et al. (42), 2009).

3.3 Data

The M_w 6.3 L'Aquila mainshock has been recorded by a dense network of receivers (e.g., Akinci et al. (5), 2010). Some of them belong to the Italian National Seismic

Network (INSN)[1], that is managed by Istituto Nazionale di Geofisica e Vulcanologia (INGV) and consists of more than 250 digital permanent seismic stations with various characteristics (Amato & Mele (7), 2008). Moreover, there are receivers of the Mediterranean Very Broadband Seismographic Network MedNet[2], that consists of 22 very broadband stations in the Euro-Mediterranean area, 13 of which are located in Italy (Mazza et al. (138), 2008). In addition, the mainshock has been recorded by strong-motion stations of the Italian Strong Motion Network (*Rete Accelerometrica Nazionale*, RAN)[3], that is managed by the Italian Civil Protection (DPC) and consists of 464 digital accelerometric receivers[4].

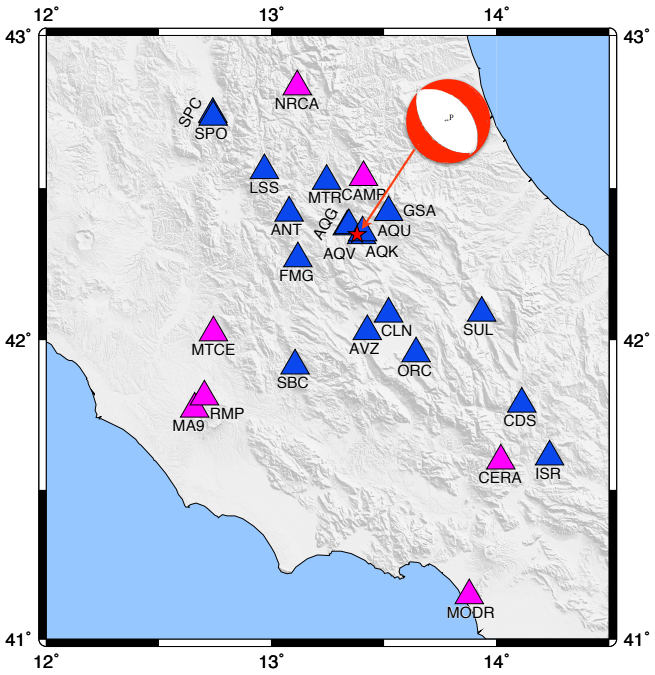


Figure 3.3: Recording stations - For the study of L'Aquila event we consider 7 velocimetric stations of INSN, indicated by violet triangles, and 18 accelerometers (17 of the RAN and one placed at the MedNet station AQU), indicated by blue triangles. The red star represents the L'Aquila mainshock.

[1] http://iside.rm.ingv.it/iside/standard/index.jsp
[2] http://mednet.rm.ingv.it/data.php
[3] Data available in the ITACA database http://itaca.mi.ingv.it/ItacaNet/; Luzi et al. (122), 2008.
[4] http://www.protezionecivile.gov.it/jcms/it/ran.wp

Among the available stations we selected only those inside the simulated volume in central Italy (section 3.2). Moreover, we rejected the velocimetric stations that showed signals saturated due to the immediate vicinity of the mainshock. The final dataset consists of 7 velocimetric stations of INSN and 17 accelerometric stations of the RAN. In addition, AQU, one of the very broadband stations of MedNet operated by INGV, was equipped by an accelerometer; thus, we include in the data the on-scale accelerometric signal of this station. The total 25 receivers are shown in Figure 3.3, together with the location of the L'Aquila mainshock.

Including accelerometric stations in the data is very useful when dealing with moderate and large earthquakes, as is the case of L'Aquila mainshock. In fact, these receivers provide on-scale seismic signals recorded very close to the event, and in other studies they have been used to estimate source parameters, seismic attenuation and local site amplification effects (e.g., Cirella et al. (42), 2009; Bindi et al. (16), 2009; Akinci et al. (5), 2010).

It is worth noting that 13 of the considered 25 stations have been also used in the source inversion of Cirella et al. (42) (2009). They are 12 accelerometers of the RAN and the accelerometer placed at the MedNet station AQU. As a consequence, a preliminary analysis has been performed in our work to compare Cirella et al. (42) (2009)'s synthetic waveforms and those simulated using SPECFEM3D (section 3.6.1). This allows us to assess if the source model implemented in our code is consistent with that obtained by Cirella et al. (42) (2009).

3.4 SEM simulation set up

As described in section 2.4, in order to perform wavefield simulations using the spectral-element code SPECFEM3D, one needs to define a model for the considered structure and seismic source. Moreover, a mesh that discretizes the simulation volume has to be constructed. The following sections present the velocity models (1D and 3D) used for central Italy (section 3.4.1), the corresponding meshes (section 3.4.2), and the L'Aquila event source representation (section 3.4.3). Moreover, we highlight the computational costs associated to both 1D and 3D simulations (section 3.4.4).

3.4.1 1D and 3D velocity models

Our objective is to accurately simulate the seismic waveforms generated by the 2009 L'Aquila mainshock in the central Italy region. The description of the source, that we use, relies on the model obtained by Cirella et al. (42) (2009) from a joint inversion of seismological and GPS data (see section 3.4.3). In that work, they used a 1D velocity profile to describe the crustal structure in central Italy. As a consequence, we are firstly interested in studying the L'Aquila earthquake wavefield generated using 1D models. Next, in order to analyse the mainshock for a more realistic structure, we construct a 3D velocity model to be used in the simulations. Both 1D and 3D models cover the considered domain, that extends 200 km x 200 km x 60 km in central Italy (see section 3.2 and Fig. 3.1 (a)).

Concerning the one-dimensional description of the structure, in this work we tested two distinct 1D models for central Italy.

The first profile substantially reproduces the crustal model used by Cirella et al. (42) (2009). This will allow us to assess if the source description has been properly implemented in our code (section 3.6.1). Cirella et al. (42) (2009) used several information to constrain the crustal structure in the region. In particular, their 1D velocity model incorporates, in depth, the 1D regional profile of Bagh et al. (11) (2007), based upon P-wave arrival times, and then a shallow low velocity layer (0-1.5 km depth), based on the model of Herrmann & Malagnini (81) (2009) from surface wave dispersion and consistent with well data. Moreover, for stations AQU and AQG they adopted the specific 1D velocity models obtained from receiver functions (see Cirella et al. (42), 2009). In our work, we consider this model of Bagh et al. (11) (2007) modified to account for the shallow low velocity layer, and it is represented in Figure 3.4 (v_p and v_s profiles labeled by "LVZ"). The profiles from the receiver functions, instead, have not been incorporated.

It is worth noting that our first 1D model (substantially equal to that in Cirella et al. (42) (2009)), described above, does not include surface topography and attenuation. These two sources of complexity, instead, have been incorporated in the second 1D structure model that we constructed for the study of L'Aquila earthquake. This allows us to reduce possible amplification effects due to neglecting the attenuation, and, in particular, it enables us to highlight the effects of the topography on wave propagation

(section 3.6.2). To construct this second 1D model, the velocity profile is again described by Bagh et al. (11) (2007)'s modified crustal model used in Cirella et al. (42) (2009) (Fig. 3.4). Next, the surface topography is imported by using Shuttle Radar Topographic Mission (SRTM) data and modifying the mesh that discretizes the simulated volume (section 3.4.2). Concerning the attenuation, it can be implemented using the code SPECFEM3D, as described in section 2.4. Practically, we estimate the quality factor Q as a function of v_s, using the scaling rule in Olsen et al. (151) (2003), that relates Q and v_s by

$$Q = 0.02\, v_s\,. \tag{3.1}$$

In 3.1 v_s is in m/s, and for the constructed 1D model each layer has a constant value of Q. This can be in the range [40-150], with step of 10, or, for the layer under the Moho it has been fixed to a very high value ($Q = 9000$), that represents a bedrock without attenuation (see also section 3.4.4).

Note that a definition of the mass density values is required by the code in order to run the simulations. Thus, in both the above described 1D models we use a 1D profile for ρ (with 6 layers, as for v_p and v_s) provided by Cirella et al. (42) (2009) (Fig. 3.4).

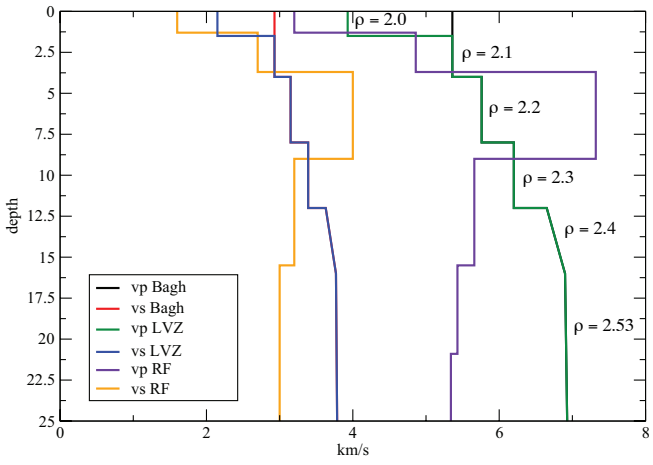


Figure 3.4: 1D velocity and density profiles - The profiles labeled by "LVZ" are those used for our 1D velocity models. The density values ρ (in g/cm^3) are shown for each layer. The label "Bagh" indicates Bagh et al. (11) (2007)'s crustal model, "RF" refers to the receiver function profiles not used in our work. Courtesy of Cirella et al. (42) (2009).

Central Italy features a complex tectonic structure (section 3.2), thus a simple 1D model can not account for all the existing lateral heterogeneities and 3D complexities. Moreover, our spectral-element code is particularly suitable for modeling seismic wave propagation in complicated realistic structures (section 2.4). Thus, aiming to reproduce L'Aquila event wavefield as accurately as possible, a 3D model for the region has been also constructed in this work.

We seek to incorporate all the available information about central Italy structure. As a consequence, we started by considering the 3D crustal tomographic model obtained by Chiarabba et al. (39) (2010), that was the most recent study in 3D for this region when we started our work. They inverted body wave arrival times from local earthquakes (occurred in 2003 and 2004) to infer v_p and v_p/v_s models down to 15-18 km depth. In addition, to support the reliability of tomographic results and to reconstruct the entire crustal structure, they used receiver functions of teleseismic waveforms, that provide constraints also on the v_s model. The structure retrieved by Chiarabba et al. (39) (2010) has strong lateral and vertical heterogeneities. Remarkable features are, for example, low velocity areas in correspondence and beneath (down to about 8-12 km depth) the main sedimentary basins, such as Fucino and L'Aquila basins, and high velocity bodies, between about 8 and 16 km depth, beneath the Simbruni Mts and the Maiella-Gran Sasso thrusts.

The tomography of Chiarabba et al. (39) (2010) is the first component of our 3D model for central Italy. However, although Chiarabba et al. (39) (2010)'s model already evidenced shallow low velocity areas in correspondence of the basins, in order to more accurately describe these features, we also consider the v_s30 geotechnical layer from Michelini et al. (140) (2008). This model defines the v_s velocities in the first 30 m depth, and it is superimposed on the 3D tomography of Chiarabba et al. (39) (2010) in order to construct our 3D model for central Italy, represented in Figure 3.5. Practically, a linear interpolation has been performed between the v_s30 model and Chiarabba et al. (39) (2010)' s velocities, also accounting for the 1D profile in Cirella et al. (42) (2009) to avoid inconsistent velocity values. Moreover, on the surface the v_p/v_s ratio has been fixed at 1.79. Based upon the mesh used for L'Aquila simulations with this 3D model (section 3.4.2), the minimum values of v_p and v_s have been set to

$$v_p^{min} = 1000.0 \text{ m/s} \qquad v_s^{min} = 600.0 \text{ m/s}\,. \tag{3.2}$$

The 3D model constructed for L'Aquila simulations, as well as the 1D model with topography, also include a description of the region below the Moho discontinuity. In particular, for both models, under the Moho depth (defined for central Italy by Di Stefano et al. (60) (2011)) we consider a homogeneous layer with velocities v_p=8000.0 m/s and v_s=4444.4 m/s (by fixing v_p/v_s=1.8; Di Stefano et al. (60), 2011). Moreover, only for the 3D case, the Moho surface is directly incorporated in the constructed mesh (section 3.4.2), based upon its topographic model given by Di Stefano et al. (60) (2011).

Concerning the density ρ, for the 3D model its values are estimated from v_p using the quadratic relationship

$$\rho \;=\; 0.025\, v_p^2 - 0.055\, v_p + 2.134\,, \tag{3.3}$$

where ρ is in g/cm^3 and v_p in km/s. Equation 3.3 has been empirically derived (Magnoni et al. (126), in prep.) by interpolating gravity data and geological models published for the Central Apennines by Di Luzio et al. (57) (2009) (see figures 7 (c) and 8 (b) therein).

Finally, surface topography and attenuation have been implemented in the 3D model as described above for the 1D profile with topography. In the case of the 3D model, however, each point of the considered volume (above the Moho) has a different value of Q, since v_s is described by a tomographic structure and not by a layered model.

3.4.2 Meshes

As previously discussed, we are interested in studying wave propagation for the 2009 L'Aquila earthquake both on 1D and 3D models. In order to implement these velocity structures in our spectral-element code, it is necessary to construct a mesh that discretizes the considered volume and its material properties (section 2.4). The mesh is composed of hexahedral elements and a value for every model parameter is assigned at each node of these elements, i.e., at each Gauss-Lobatto-Legendre (GLL) point.

In order to run the simulations with the 1D models in section 3.4.1, a first mesh has been constructed using the mesh generator GEOCUBIT (section 2.4). This mesh is very simple and completely homogeneous, i.e., it does not incorporate any discontinuity. It covers the considered domain of 200 km x 200 km x 60 km in central Italy (section 3.2), and it consists of 259'200 hexahedral elements with a size of 1 km. When used to implement the 1D model without topography, the mesh is completely flat on

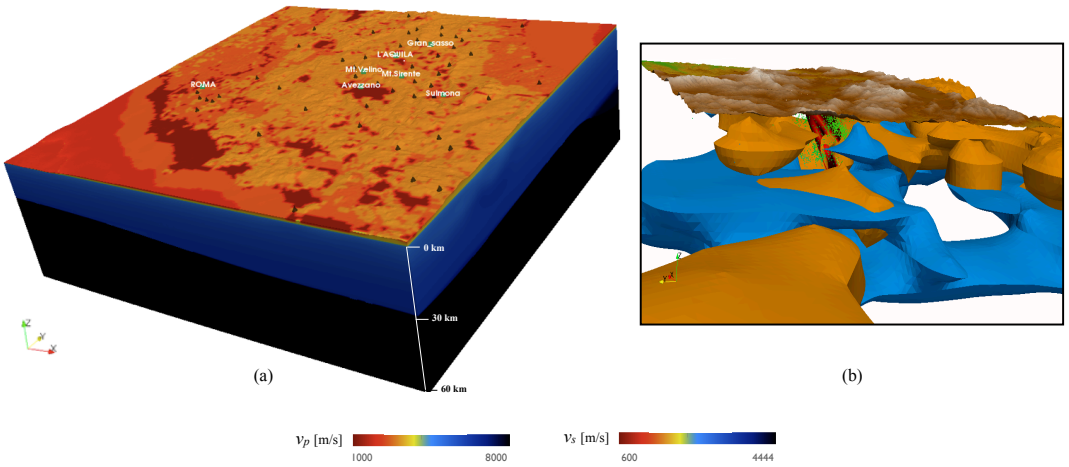


Figure 3.5: 3D velocity model - (a) View from SE of the considered 3D velocity model. v_p and v_s scales are at bottom. The volume is 200 km x 200 km horizontally and 60 km deep, and the black layer is the fast region under the Moho. (b) Close-up of the epicentral area viewed from NW (as in Fig. 3.2 (a)). Blue features have v_p=6400 m/s, orange features have v_p/v_s=1.84 (Chiarabba et al. (39), 2010).

top. On the contrary, considering the 1D model with topography, the Shuttle Radar Topographic Mission (SRTM) data are used to model the surface of the mesh. The resulting topographic profile, that has an original resolution of 90 m (Jarvis et al. (90), 2008), is then smoothed at 1 km. Moreover, in this case the GLL points in the layer below the Moho have the values of v_p and v_s given by Di Stefano et al. (60) (2011) (section 3.4.1). The numerical resolution of the simple mesh for 1D models is ∼0.5 Hz.

To implement the 3D model for central Italy in section 3.4.1, a more complex mesh is necessary. Thus, using GEOCUBIT, a second mesh has been constructed, again covering the 200 km x 200 km x 60 km volume in central Italy.

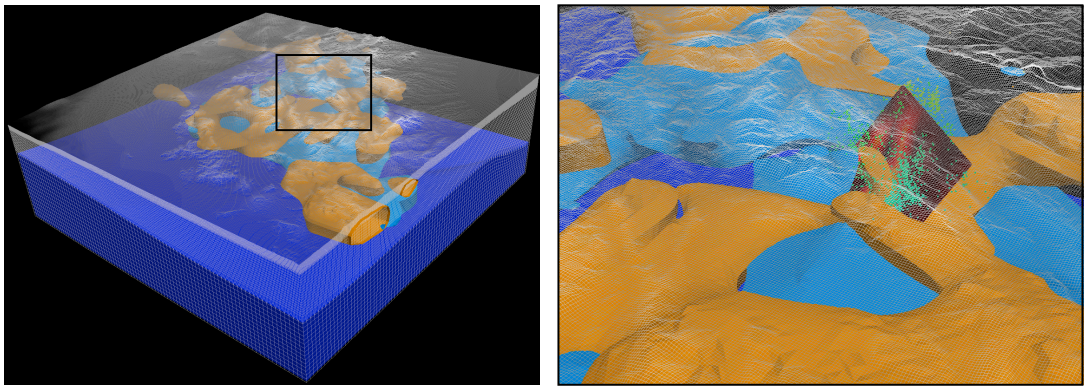
This mesh, represented in Figure 3.6, is conforming and contains 7.6 million of hexahedral elements, that have a dimension of 200 m on the surface. The size of the elements is then triplicated along depth, in order to allow for velocity increase. As a consequence of the small element size, the mesh has a very high numerical resolution, up to about 5 Hz. Thus, in principle, it allows us to accurately simulate seismic wave propagation at these high frequencies. For L'Aquila simulations, as inferred from the results of our study (section 3.7.1), the structure and source models do not yet reach this level of accuracy, limiting the maximum affordable frequency to lower values (about 0.5

Hz). However, this mesh will become particularly useful when more accurate models will be sought (e.g., after several iterations of the adjoint inversion in chapter 4).

The total number of GLL points is 49.8 x 10^7, and to each of them is assigned a value of v_p and v_s, by interpolating the considered 3D heterogeneous model (using a suitable routine that we implemented in the code). Then, based upon relations 3.1 and 3.3 each point has also a value of Q and ρ, respectively.

The mesh for the 3D model honors the free surface topography, implemented using the SRTM profile smoothed at 200 m. Moreover, it honors the surface of the Moho discontinuity, whose topography is provided by Di Stefano et al. (60) (2011). Below this discontinuity, the values of the model parameters assigned to each GLL point are constant, describing a high velocity homogeneous layer (section 3.4.1).

Such a complex mesh requires about 12 hours to be constructed using 308 cores of the cluster available at INGV, and the corresponding wavefield simulations based upon SPECFEM3D are highly expensive (see section 3.4.4 for more details).



(a) (b)

Figure 3.6: Mesh for L'Aquila simulations - (a) View from SE of the mesh used to implement the 3D velocity model of central Italy. The discretized volume is 200 km x 200 km x 60 km, blue features have v_p=6400 m/s, orange features v_p/v_s=1.84 and the dark blue layer in depth is under the Moho surface. One can note the element size increasing from the top to the bottom. (b) Close-up of the black inset in fig.(a). The fault plane of the L'Aquila event is also shown, together with the aftershocks that occurred in the 10 days following the mainshock (green squares).

3.4.3 Source model

Our representation of the source for the 2009 L'Aquila event is based on Cirella et al. (42) (2009)'s model, that includes primarily seismological strong motion data in the inversion. Based upon that model, the source is described by a finite fault with strike and dip angles fixed at N133°E and 54° to the SW, respectively. The considered fault plane is 28 km long and 17.5 km wide, and to perform the inversion it is subdivided into subfaults of 3.5 km x 3.5 km. Their model is described in terms of the corners of each patch, at which the model parameters - peak slip velocity, slip direction, rupture time and rise time - are fixed. Within each subfault, instead, these parameters can vary by interpolating the nodal values. The corners are initially 54, but during the inversion procedure Cirella et al. (42) (2009) perform a further interpolation with ∼258 m resolution, in order to make the inversion accurate up to about 0.5 Hz. As source time function they use a regularized Yoffe function (Tinti et al. (195), 2005) with time to peak slip velocity T_{acc}=0.225 s. Moreover, they use the 1D crustal model described in section 3.4.1. Their nonlinear inversion technique results in an average stable model of the earthquake rupture parameters in agreement with the data, from which the final slip distribution is derived (see Fig. 3.2).

In order to implement Cirella et al. (42) (2009)'s source model in our code, we consider a finite fault plane with the same dimensions and the same strike and dip angles. Moreover, the fault is again subdivided into subpatches with a resolution of 258 m. However, in our code a finite source is represented as a superposition of point sources, each of which is described by the 6 moment tensor components, location, rise time and rupture time. Thus, in our case we locate the point sources at the center of each 258 m x 258 m patch (moving the control points from the corners considered by Cirella et al. (42) to the centers), in order to reproduce exactly the spatial extension of Cirella et al. (42) (2009)'s fault. This results in a finite fault discretized into 7063 point sources, each of which has source parameters obtained by interpolating those of Cirella et al. (42) (2009)'s inverted model.

More practically, to use the above fault model in SPECFEM3D, every point source is described by a Gaussian source time function (section 2.4), with a half duration that is a half of the corresponding rise time, calculated by interpolating the values given by

Cirella et al. (42) (2009)[1]. The rupture time of each point source in a patch center is the minimum rupture time among those given by Cirella et al. (42) (2009) at the 4 patch corners. Moreover, to assign the moment tensor to each point source, one needs the scalar seismic moment M_0 and the source mechanism (Aki & Richards (3), 1980). Strike and dip are fixed (see above), and the rake is interpolated from Cirella et al. (42) (2009)'s values, as the slip required to calculate M_0. The area of the patch is given by the resolution. The remaining parameter to be estimated is the shear modulus μ for every source (required for M_0). This parameter is calculated from v_s and ρ, and in this work we are interested in studying L'Aquila event for different structure models, both 1D and 3D (section 3.4.1). Thus, to obtain a finite source model for our simulations consistent with the considered velocity model, we assign the values of μ of the subsources based upon v_s and ρ of each structure description. As a consequence, two different finite source models result:

1. A finite source model discretized into 7063 point sources (258 m resolution) with moment tensors based upon the 1D model for central Italy (regardless of accounting for topography; section 3.4.1);

2. A finite source model discretized into 7063 point sources (258 m resolution) with moment tensors based upon the 3D model for central Italy (section 3.4.1).

Note that, for the fault model in 2., although the 3D structure is accounted for in moment tensor calculations, all the other source parameters are derived from Cirella et al. (42) (2009), i.e., from an inversion based on a 1D velocity model (as in case 1.). Thus, the results of simulations for the 3D model will be affected by an intrinsic incompatibility between the considered structure model and the source description. This represents a limit of our work, highlighting the necessity for source models derived from inversions in 3D structures (e.g., Liu et al. (119), 2004; Kim et al. (97), 2011; see sections 3.7 and 3.9).

Finally, it is worth noting that, always starting from Cirella et al. (42) (2009)'s model, other source interpolations have been tested, also with a resolution higher than that in Cirella et al. (42) (2009) (i.e., a patch size <258 m). This may increase the

[1]Practically, the rise time for each point source in the center of a patch is calculated from the values given by Cirella et al. (42) (2009) at the corners as the difference between the time at which the corner, that moves last, ends moving, and the time at which the corner, that moves first, starts moving.

resulting frequency of the source model. However, pushing the resolution at values too high compared to frequencies resolved by Cirella et al. (42) (2009), the reliability of inferred source parameters is affected. Thus, we prefer to reproduce Cirella et al. (42) (2009)'s model as accurately as possible always using a resolution of 258 m, and we obtain slip and rise time on-fault distributions that agree very well with those deduced by Cirella et al. (42) (2009) and represented in Figure 3.2 (b)-(c).

3.4.4 SEM simulations and computational requirements

The structure and source models, together with the corresponding meshes, described in previous sections, represent the key ingredients to run the simulations for L'Aquila event using SPECFEM3D.

For each of the three considered velocity models, described by the suitable mesh (sections 3.4.1 - 3.4.2), we run a simulation with the specific finite source model, whose values of M_0 rely on the given structure (model 1. or 2. in section 3.4.3).

As presented in section 2.4, the code allows one to implement both anisotropy and attenuation. For all our simulations anisotropy has been neglected. Concerning attenuation, it can be included in different ways. One option is to directly assign a specific value of Q for each GLL point (assigning it in the mesh). Otherwise, one can use Olsen et al. (151) (2003)'s model and define Q as a function of v_s. In this case, one refers to *simple Olsen attenuation*, if the values of Q are calculated using equation 3.1, when v_s<2000 m/s, and a similar relation with a factor 0.1 instead of 0.02, when v_s>2000 m/s. Instead, one refers to *discrete Olsen attenuation*, if Q is always calculated using 3.1 and, based upon its integer value, it is forced to assume a discrete value within [40-150], with step of 10. We tested all these three possibilities using the 3D model for L'Aquila earthquake [1]. We found that using the discrete Olsen attenuation reduces observed amplification effects with respect to data for most of the analysed synthetic timeseries. Thus, we choose to consider this implementation of the attenuation in the simulations for the 3D model and the 1D model with topography.

To summarize, we ran three different simulations for our study of the 2009 L'Aquila mainshock:

[1] In the case of Q values defined in the mesh we assume a constant value for all GLL points due to the lack of a detailed model of Q.

1. A simulation for the 1D velocity model without topography, implemented in the simple homogeneous flat mesh, and the corresponding finite source model of L'Aquila event (model 1. in section 3.4.3). The attenuation is not implemented.

2. A simulation for the 1D velocity model with topography, implemented in the simple homogeneous mesh with topography, and the corresponding finite source model of L'Aquila event (model 1. in section 3.4.3). Implementation of the discrete Olsen attenuation.

3. A simulation for the 3D velocity model, implemented in the complex mesh, and the corresponding finite source model of L'Aquila event (model 2. in section 3.4.3). Implementation of the discrete Olsen attenuation.

Each simulation produces the forward synthetic wavefield at all the 25 stations considered in this work (section 3.3). The L'Aquila event has a moderate size and it ruptured relatively quickly (e.g., Cirella et al. (42), 2009; Çelebi et al. (28), 2010). Thus, considering the distance of the receivers from the mainshock, for every station we simulate timeseries of 60 seconds, that in general include both body and surface waves. The considered timestep is $1x10^{-3}$ s (i.e., the sampling rate of the synthetics is originally 10^3 Hz).

To perform all the simulations, we exploited the High-Performance Computing (HPC) resources of INGV[1]. In particular, we ran parallel jobs on the cluster *ELIOS*, that has 64 compute nodes, each with 2 quad-core AMD Opteron 2374 processors at 2.4 GHz and with 16 GB RAM (512 total cores, 2 GB RAM/core).

The computational times required for our simulations using *ELIOS* are summarized in what follows:

- Simulations of 60 s, based upon the finite source model and the 1D velocity models (with and without topography), require 45 minutes on 32 cores;

- Simulations of 60 s, based upon the finite source model and the 3D velocity model, require 32 hours on 308 cores.

The very different simulation times are due to the different meshes adopted. On the contrary, the number of point sources, that discretize the finite source, and the number of considered stations do not affect the computational requirements.

[1] http://hpc.rm.ingv.it/

3.5 Processing of the dataset

Once the synthetic seismograms have been obtained using SPECFEM3D (section 3.4.4), they have to be compared with the observed waveforms in order to infer the results of our simulations.

Data and synthetics need to be processed as closely as possible, in order to avoid discrepancies between the two timeseries originated just by the different processing. As previously described (section 3.3), we use signals from both velocimetric and accelerometric stations. Thus, two distinct procedures have been applied to process these waveforms.

Concerning the accelerometric data, we use uncorrected strong motion timeseries only from digital RAN instruments. These data are available from the Italian strong motion database ITACA (Luzi et al. (122), 2008) in a pre-processed form, in order to remove the so called non-standard errors (Massa et al. (136), 2009). Moreover, they do not need to be corrected for the instrument response, since the response is flat up to 50 Hz (Massa et al. (136), 2009). Thus, these data, together with SEM synthetics in acceleration[1], are processed in this work using the following procedure:

1. Resample data and synthetics at 200 Hz;
2. Integrate data and synthetics to obtain velocity timeseries;
3. Low-pass filter data and synthetics with a corner frequency of 0.5 Hz and a two-pole, two-pass Butterworth filter;
4. High-pass filter data and synthetics with a corner frequency of 0.02 Hz and a two-pole, two-pass Butterworth filter;
5. Multiply synthetics by 100 to obtain velocities in cm/s.

The same processing has been applied to the timeseries recorded by the accelerometer located at station AQU (section 3.3), except for the fact that its observed data were initially in m/s^2, thus they are converted to cm/s^2.

[1]SPECFEM3D produces, as output of the simulations, synthetic seismograms in displacement, velocity and acceleration; depending on the processing we need to perform, one of them is used.

Concerning the velocimetric recordings, a correction for instrumental response needs to be applied. Several processing tests have been performed on the considered waveforms, and it results that, in the range 0.02-0.5 Hz, spurious effects on the observed data often arise when they are deconvolved with instrument response. Thus, we preferred not to affect the data, but to convolve the synthetics with the instrumental responses. The processing applied to the velocity timeseries, both observed and synthetic, is outlined in what follows:

1. Resample data and synthetics at 200 Hz;
2. Convolve the synthetics with the corresponding instrumental responses, while keeping them in velocity (in count/s);
3. Low-pass filter data and synthetics with a corner frequency of 0.5 Hz and a two-pole, two-pass Butterworth filter;
4. High-pass filter data and synthetics with a corner frequency of 0.02 Hz and a two-pole, two-pass Butterworth filter.

The choice of the frequency range to filter data and synthetics is primarily based on the fact that 0.02-0.5 Hz is the frequency band considered in the source inversion of Cirella et al. (42) (2009). The L'Aquila event source model, that we use, is derived from that of Cirella et al. (42) (2009) (section 3.4.3). Thus, it is reasonable that the accuracy of the problem may be limited within Cirella et al. (42) (2009)'s frequency range, although the numerical resolution only based upon the mesh could be much more higher (section 3.4.2).

As introduced in section 3.3, a useful test is to compare our results with those of Cirella et al. (42) (2009), and we will discuss this analysis in section 3.6.1. To this end, we use the synthetic timeseries of Cirella et al. (42) (2009) for the accelerometric stations we have in common. These synthetics underwent the following processing:

1. Sampling at 200 Hz;
2. Integration to obtain velocity timeseries (in cm/s) (by Cirella et al. (42), 2009);
3. Filtering between 0.02-0.5 Hz with a two-pole, two-pass Butterworth filter (by Cirella et al. (42), 2009).

It is worth noting that Cirella et al. (42) (2009)'s synthetics have down-oriented vertical components, thus, in order to be consistent with their results, the Z components of all our synthetics and data have been multiplied by -1. Moreover, timeseries of 60 seconds are always considered.

3.6 Simulations with 1D velocity models: topography effects

We start the analysis of L'Aquila mainshock by considering the wavefield that propagates in 1D models of central Italy structure. The 1D profiles that we consider (section 3.4.1) differ from each other mainly for neglecting or including the surface topography. Firstly, the 1D model without topography will offer us the possibility of comparing our results with those of Cirella et al. (42) (2009) (section 3.6.1). Moreover, a comparison of the synthetics of both our 1D models with respect to the observed data will allow us to highlight possible effects related to topographic features (section 3.6.2).

3.6.1 Comparison with Cirella *et al.* (2009)

The source model adopted in this work for the 2009 L'Aquila event is derived from that of Cirella et al. (42) (2009) (section 3.4.3). Their inversion is based upon a 1D crustal profile for central Italy, that does not include the topography of the free surface. As a consequence, a preliminary useful analysis should consist in verifying if we are able to reproduce the behaviour of the synthetic seismograms obtained by Cirella et al. (42) (2009), assuring us that the source model has been properly described. This is the primary reason why we constructed a first 1D model of the region substantially equal to that of Cirella et al. (42) (2009) (section 3.4.1).

As explained in section 3.4.4, in order to perform this preliminary study, we run a simulation for the 1D velocity model without topography and the corresponding source model for L'Aquila earthquake (simulation 1. section 3.4.4). Then, we analyse the behaviour of our synthetics and those of Cirella et al. (42) (2009) with respect to data. Figure 3.7 shows the seismograms of Cirella et al. (42) (2009) (first column, red) and our SEM seismograms (second column, red) compared to the data (always black) for three stations that we have in common with Cirella et al. (42) (2009). Moreover, it also shows for these receivers the synthetics obtained from our simulation with the

1D model with topography (simulation 2. section 3.4.4), to have a first indication of topographic effects (third column, red).

Looking at the synthetics in Figure 3.7, one initially notes that the first part of the seismograms is only slightly influenced by the considered structure model. In particular, the behaviour in the first seconds of the timeseries seems to be primarily related to the source description (based upon Cirella et al. (42) (2009)'s inversion), that is common in all cases. In these initial portions it is difficult to distinguish the effect of the topography, although its introduction causes weak changes that may improve the fit also there. The topographic effect, instead, is dominant in the coda of the seismograms (comparing the results in the first two columns of Fig. 3.7 with those in the last column), allowing one to model many observed features, otherwise missed. This will be discussed in more details in the next section 3.6.2.

Now, we are particularly interested in comparing our synthetics for the flat 1D model to those of Cirella et al. (42) (2009) (red in 2^{nd} and 1^{st} columns of Fig. 3.7, respectively). The L'Aquila source description is almost the same in the two cases (section 3.4.3), as well as the velocity profile (section 3.4.1). Thus, one could expect an almost perfect match between the corresponding synthetics. However, few differences are evident, although they are very slight compared to the changes in the synthetics due to the introduction of the topography (3^{rd} column of Fig. 3.7). Regarding station AQU, the differences can be partially explained by considering that Cirella et al. (42) (2009) adopted the specific velocity model retrieved from receiver functions to compute the synthetics for this station (the same is for station AQG, not shown in the picture). In our case, instead, we use the same crustal profile for the entire domain.

A more general cause of the discrepancy between Cirella et al. (42) (2009)'s synthetics and those for our 1D flat model may be attributed to the difference between the considered source time functions. Cirella et al. (41) (2006) showed the effect of different source time functions (Fig. 3.8 (a)) on the ground motion computed for the 2000 Western Tottori earthquake (for frequencies up to 1 Hz). In particular, they observed the results at the nearest and the furthest stations from the event (Fig. 3.8 (b) and (c), respectively). They concluded that the calculated ground motion depends on the choice of the source time function, especially for stations next to the source, whereas differences decrease moving farther from the causative fault. Concerning L'Aquila event, as presented in section 3.4.3, Cirella et al. (42) (2009) used, at each corner point, a

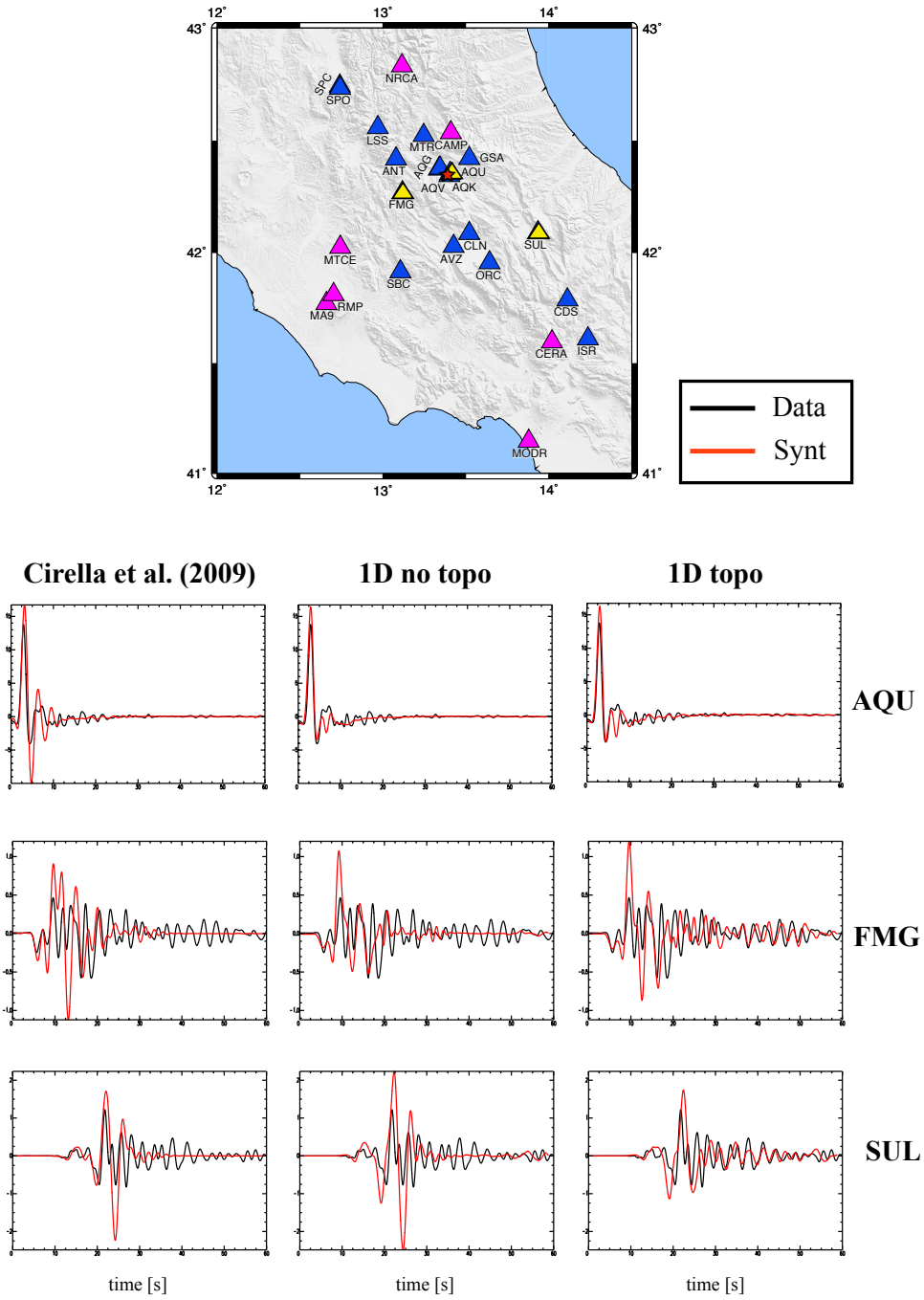


Figure 3.7: Comparison with Cirella et al. (2009) - Z-down component velocity (in cm/s) for the 3 stations represented by yellow triangles in the map. The red star in the map is the L'Aquila event. The recorded timeseries are black, the synthetic timeseries are red and have been obtained by Cirella et al. (42) (2009) (1^{st} column), and by SEM simulations using the 1D velocity model without topography (2^{nd} column) or the 1D velocity model with topography (3^{rd} column). The seismograms are filtered within 0.02-0.5 Hz and the y axis is the same for each station.

source time function represented by a regularized Yoffe with T_{acc}=0.225 s, as the violet curve in Figure 3.8 (a). On the contrary, based on our code, we attributed to each of the point sources, that discretize the finite fault, a Gaussian source time function with specific half duration. This Gaussian (Fig. 3.8 (d)) is similar to the modified cosine function (green) in Figure 3.8 (a). Comparing Cirella et al. (42) (2009)'s synthetics to those for our 1D flat model (green and red seismograms in Fig. 3.8 (e), respectively), one notes that the discrepancies are consistent with those observed by Cirella et al. (41) (2006) for the two different source time functions (Fig. 3.8 (b) and (c)). Moreover, the three considered stations have an increasing epicentral distance and the difference between the synthetics decreases moving farther from the source. Thus, based upon Cirella et al. (41) (2006)'s conclusions, we can consider the observed slight discrepancies between our (1D flat) and Cirella et al. (42) (2009)'s synthetic timeseries mainly due to the different functions used to model the source.

As a conclusion, considering Figure 3.7, one can infer that, using the same velocity model and the same source description (except for its functional form), we are able to simulate synthetics that behave, with respect to data, similarly (sometimes better, e.g., AQU and FMG Fig. 3.7) to Cirella et al. (42) (2009)'s synthetics. This suggests that we properly implemented Cirella et al. (42) (2009)'s source model, despite the different source time function that is required by our code.

The above discussion supplies a possible explanation for the observed differences between our 1D flat model simulations and Cirella et al. (42) (2009)'s results, remarking, however, that they are generally slight with respect to topographic effects. The next step will be to focus on the different synthetic ground motion with and without accounting for topography (section 3.6.2). Only the waveforms of our simulations will be analysed, assuming that our 1D flat model substantially reproduces Cirella et al. (42) (2009)'s results.

It is worth noting that further developments of the above study may include a numerical test that considers for L'Aquila earthquake a point source described by the two different source time functions (Yoffe and Gaussian). This will definitively show how the source time function influences the ground motion in our case.

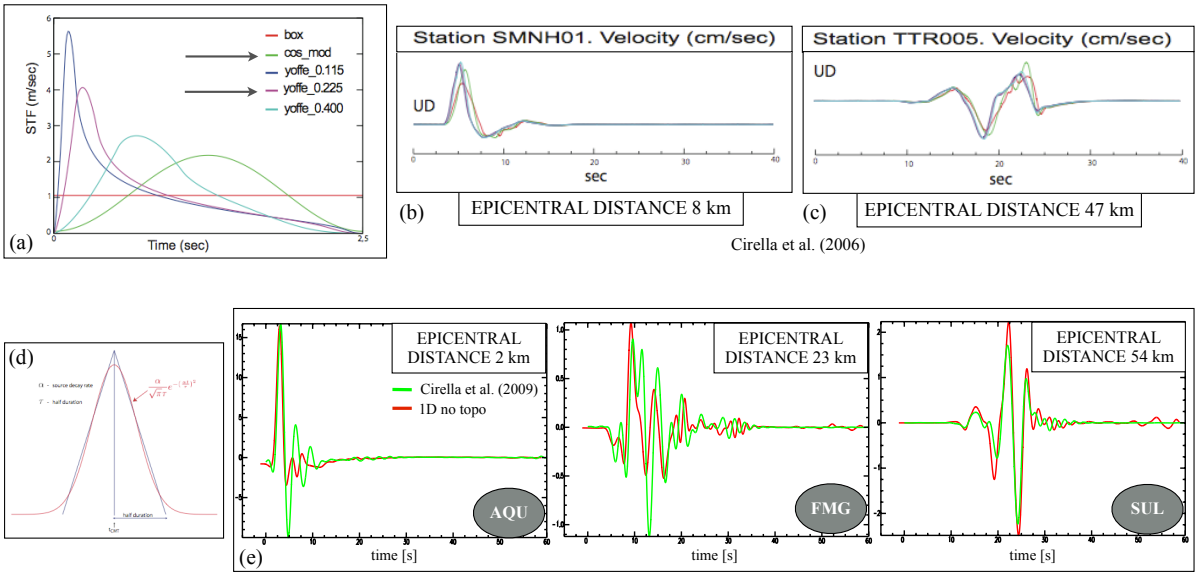


Figure 3.8: Source time function effect - (a) Source time functions (STFs) tested by Cirella et al. (41) (2006) for the 2000 Western Tottori earthquake. The number associated to each Yoffe indicates the time to peak slip velocity T_{acc} (in sec). (b)-(c) Comparison between Z-component seismograms calculated with different STFs, for the nearest (b) and the farthest (c) stations from the Tottori earthquake. Courtesy of Cirella et al. (41) (2006). (d) Gaussian source time function (red) used by our code SPECFEM3D 2.0. (e) Comparison between Cirella et al. (42) (2009)'s synthetics (green) and our synthetics for the 1D flat model (red) for three stations that recorded the 2009 L'Aquila event. The timeseries are Z-component velocities (in cm/s). See the map in Fig. 3.7 for station locations.

3.6.2 1D models with and without topography

Several studies (e.g., Komatitsch & Vilotte (104), 1998; Komatitsch et al. (109), 2004; Lee et al. (113, 114, 115), 2008, 2009a,b) enhance the importance of including a realistic description of the surface topography in the models, in order to accurately reproduce seismic wave propagation and surface ground motion. In addition, these studies highlight the capability of SEM to successfully implement such a complexity in the simulations, also at a very high resolution (e.g., Lee et al. (114), 2009a). In general, surface topography can cause complex wave propagation behaviour, with seismic energy reflected and scattered by the mountains (as shown by several studies, see, e.g., Massa et al. (137) (2010) and reference therein). The ground motion, due to topographic effects, is mainly amplified on the tops of hills and ridges (e.g., Çelebi (27) 1987; Kawase & Aki (95), 1990; Spudich et al. (180), 1996; Bouchon & Barker (19), 1996; Bouchon et al. (20), 1996; Komatitsch & Vilotte (104), 1998; Lee et al. (114), 2009a), and multiple reflections may prolong the seismic shaking (e.g., Lee et al. (113), 2008). Remarkable variations of the peak ground velocity and acceleration have been also attributed to topography, highlighting its role in seismic hazard assessment (e.g., Komatitsch & Vilotte (104), 1998; Lee et al. (114, 115), 2009a,b; see section 3.8).

In central Italy, amplifications of the ground motion related to topography have been also observed (e.g., Marra et al. (132), 2000; Donati et al. (63), 2001; Massa et al. (137), 2010; Marzorati et al. (135), 2009; Pischiutta et al. (161), 2010). In general, relevant topographic reliefs are present in the simulation volume considered in our study (section 3.2). This suggests that incorporating topography in our models is fundamental to properly simulate ground motion. Seismogram comparisons in Figure 3.7 provide a first indication of the valuable information for seismic wave modeling introduced by considering the topography. To study in more details these effects, in this section we analyse the synthetics simulated, for all the considered stations, using the 1D velocity models with and without topography (simulations 1. and 2. section 3.4.4, respectively).

Figures 3.9 - 3.10 show the behaviour, with respect to the data (black), of the synthetics for the 1D model with topography (red) against the corresponding synthetics for the 1D flat model of central Italy (green). The timeseries are filtered between 0.02-0.5 Hz and all the three ground velocity components are shown. In general, as

already highlighted in Figure 3.7, the first part of the seismograms seems only slightly influenced by the introduction of topography, since very weak variations are evident between the synthetics for the two models. This suggests again that the first seconds of the waveforms could be mainly related to source effects. On the contrary, the principal evidence in all the examples is that, considering the 1D model with topography, instead of the flat one, we are able to well reproduce many features in the coda of the seismograms. This is often evident on all ground motion components (e.g., stations SBC Fig. 3.9 and NRCA Fig. 3.10), but almost always in the vertical component coda, that show examples of very good fit (e.g., FMG Fig. 3.9). Considering the station locations with respect to the mainshock, in Figure 3.10 it is also interesting to note the effect of source directivity, that for L'Aquila event has been evidenced towards SE (e.g., Akinci et al. (5), 2010). This causes larger amplitudes and higher frequency content in the seismograms (both observed and synthetics) of CERA with respect to NRCA.

The principal results observed in the above examples (Figs. 3.9 - 3.10) are valid in general for the whole dataset. Thus, many of the modulations in the seismogram coda, related to late surface wave arrivals (e.g., Clouser & Langston (45), 1995; Komatitsch & Vilotte (104), 1998), are fitted only by introducing a topographic description in the structure model. Hence, these features may be likely attributed to reflections and reverberations caused by the topographic effects of central Apennines ridges. Examples of the improved agreement are mostly evident on the Z components of the timeseries, while less on the horizontal components. This result may be explained considering the large vertical scalelength of the topographic relief in central Italy, that locally influences particularly the vertical component of ground motion.

It is worth noting that Cirella et al. (42) (2009)'s source inversion is based upon a flat 1D velocity model. As a consequence, the effects of the topography may have been partially mapped into the inverted source model. The results discussed above lead us to conclude that, although the topographic effect may be partly accounted for in the source description, an explicit implementation of the topography is necessary when we seek to accurately reproduce seismic wave propagation.

The examples in Figures 3.9 - 3.10, as well as the whole dataset, show, however, that many features in the seismograms are still not matched. This suggests that a description of the topography at a higher resolution is required, since smaller-scale heterogeneity, evident at higher frequencies, may produce even stronger effects (e.g.,

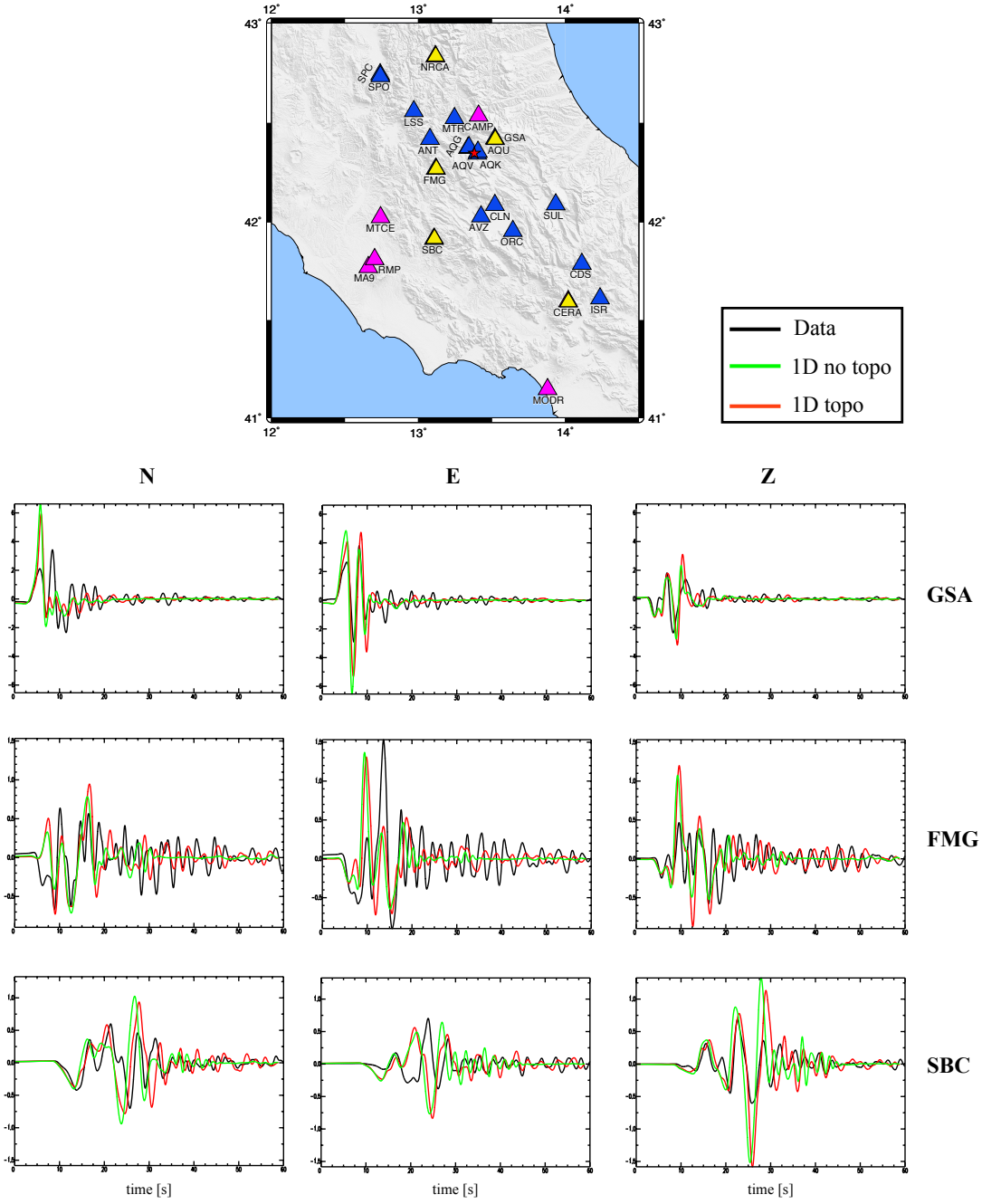


Figure 3.9: Comparison between 1D velocity models - Three-component velocity (in cm/s) for the stations written on the right and represented by yellow triangles in the map. The red star in the map is the L'Aquila event. The recorded timeseries are black, the synthetic timeseries obtained by SEM simulations using the 1D velocity model without topography are green, the SEM synthetics for the 1D model with topography are red. The seismograms are filtered within 0.02-0.5 Hz and the y axis is the same for each station.

Lee et al. (114, 115), 2009a,b). However, at present, the resolvable frequencies of our simulations are limited by the adopted source and structure models (section 3.7.1). Moreover, the remaining discrepancies between data and synthetics mainly highlight the necessity for a 3D description of the structure, that accounts for all the other complex heterogeneities in the region, besides topography. This will be the object of the analyses in section 3.7.

Finally, one should remark that another difference between the two 1D models considered in this section is that the flat one does not include wave attenuation. The implementation of this feature in the simulations is also fundamental, especially if the considered region is characterized by low wave-speed sedimentary basins (e.g., Olsen et al. (151), 2003; Komatitsch et al. (109), 2004; Lee et al. (113), 2008). A detailed description of the quality factor has not been included in our simulations, due to the lack of an accurate, comprehensive model of Q for the considered central Italy region. However, Figures 3.9 - 3.10 suggest that including the attenuation contributes to improve the fit of observed data, with respect to a completely elastic case, and it seems essential to study realistic ground motion behaviour (see also section 3.8).

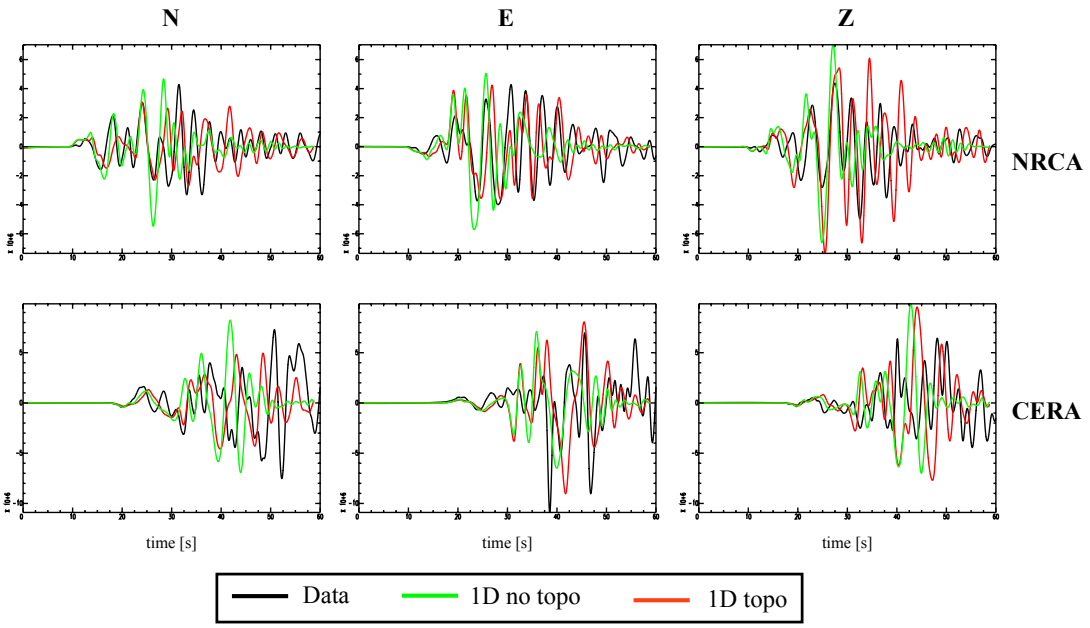


Figure 3.10: Comparison between 1D velocity models - Same as Figure 3.9, except for the fact that the velocities are in count/s.

3.7 Simulations with a 3D velocity model: complex structure effects

The complex heterogeneous structure (both in depth and horizontally) of central Italy and the results in the previous section strongly suggest the need to introduce in the simulations a 3D more thorough representation of the considered domain. Together with a detailed description of the topography, modeling lateral variations in velocity (and also density and attenuation) is of paramount importance in order to accurately and realistically reproduce the ground motion of a complex region, as revealed by many studies (e.g., Olsen et al. (151), 2003; Komatitsch et al. (109), 2004; Lee et al. (113, 114, 115), 2008, 2009a,b; Stupazzini et al. (184), 2009; Tape et al. (189), 2010). In particular, areas characterized by low wave-speed sedimentary basins may show a very different ground response, compared to rock sites, with strong amplifications, multiple reverberations and increasing signal durations (e.g., Olsen et al. (149), 2006; Lee et al. (113), 2008; Chaljub et al. (32), 2007; Rovelli et al. (170), 2001). Moreover, these effects can be intensified and complicated by the presence of topographic ridges, that produce themselves amplifications (section 3.6.2) and also, causing several reflections, force the wavefield to multiple paths inside the surrounding basins (e.g., Lee et al. (113, 115), 2008, 2009b). Significant site amplifications have been observed for the L'Aquila event in the neighbouring zones (e.g., Bindi et al. (16), 2009; Ameri et al. (8), 2009; Akinci et al. (5); 2010; Çelebi et al. (28), 2010); at L'Aquila city, for example, based on previous studies (e.g., De Luca et al. (54), 2005), these effects are related to the presence of a basin, filled of lacustrine sediments (with a maximum depth of 250 m). Several other sedimentary basins, moreover, feature the central Italy region (section 3.2), and they can have a similar influence on the ground motion (e.g., Castro et al. (25), 2004; Bindi et al. (15, 16), 2004, 2009). These effects have also a primary importance for seismic hazard assessment implications (e.g., Komatitsch et al. (109), 2004; Stupazzini et al. (184), 2009), enhancing our interest in modeling them in detail (see also section 3.8).

As introduced in section 2.2 and shown in many of the above cited studies, SEM is very well suited to accurately implement all the complexities of a realistic model. Thus, using our code and the 3D structure of central Italy (section 3.4.1), we performed a wavefield simulation with the corresponding source representation for L'Aquila event

(simulation 3. section 3.4.4). In this section we discuss the results obtained by comparing to the data the SEM synthetics for the 3D model, with respect to the case with a 1D model.

Figures 3.11 - 3.14 show the synthetic seismograms (red) for the 3D velocity structure compared to the observed timeseries (black) recorded at some of the considered stations (second row for each station). Comparisons of the data and synthetics for the 1D model with topography are also shown to highlight the variations (first row for each station). The frequency range is 0.02-0.5 Hz.

We observe, as expected, that a modulation of the seismogram coda becomes often very evident when the 3D complexities are taken into account (e.g., AQK and AQU Fig. 3.11, MA9 Fig. 3.13). In particular, an improvement of the fit, both of amplitudes and especially of phases, is shown mainly by the horizontal components of the ground motion, where in general the topographic effect is less visible (e.g., AQU Fig. 3.11, FMG Fig. 3.13, RMP Fig. 3.14). For the vertical components, instead, when the 1D model with topography already provides a good matching, the 3D model often presents excessive amplifications of the signals with respect to data (e.g., SPO Fig. 3.12, FMG Fig. 3.13). In many examples, these large amplitudes for the 3D case are observed on all the components (e.g., SPO Fig. 3.12, NRCA Fig. 3.14), whereas sometimes this model increases the fit, mostly of phases, everywhere (e.g., AQK Fig. 3.11, MA9 Fig. 3.13, RMP Fig. 3.14). An interesting case is station GSA (Fig. 3.11), for which adopting the 3D model causes a strong amplification on all the components between about 7.5 and 12.5 s. This may suggest that the considered 3D structure locates a too low wave-speed feature in proximity of this station, probably causing a spurious too strong basin effect.

The above discussed examples are substantially representative of the behaviour of the whole dataset. Thus, in general, considering a 3D description of the structure, we are able to reproduce, although often overestimated, many of the features visible in the observed seismograms for all ground motion components. This confirms that a 3D modeling of the structure is essential to capture the waveform complexities, that, as expected, are strongly related to the lateral heterogeneities in the region, such as sedimentary basins. These features seem to affect mainly the coda of horizontal component seismograms, where a good fit due to the 3D model is especially evident. On the other hand, the examples in Figures 3.11 - 3.14, as well as the whole dataset, seem

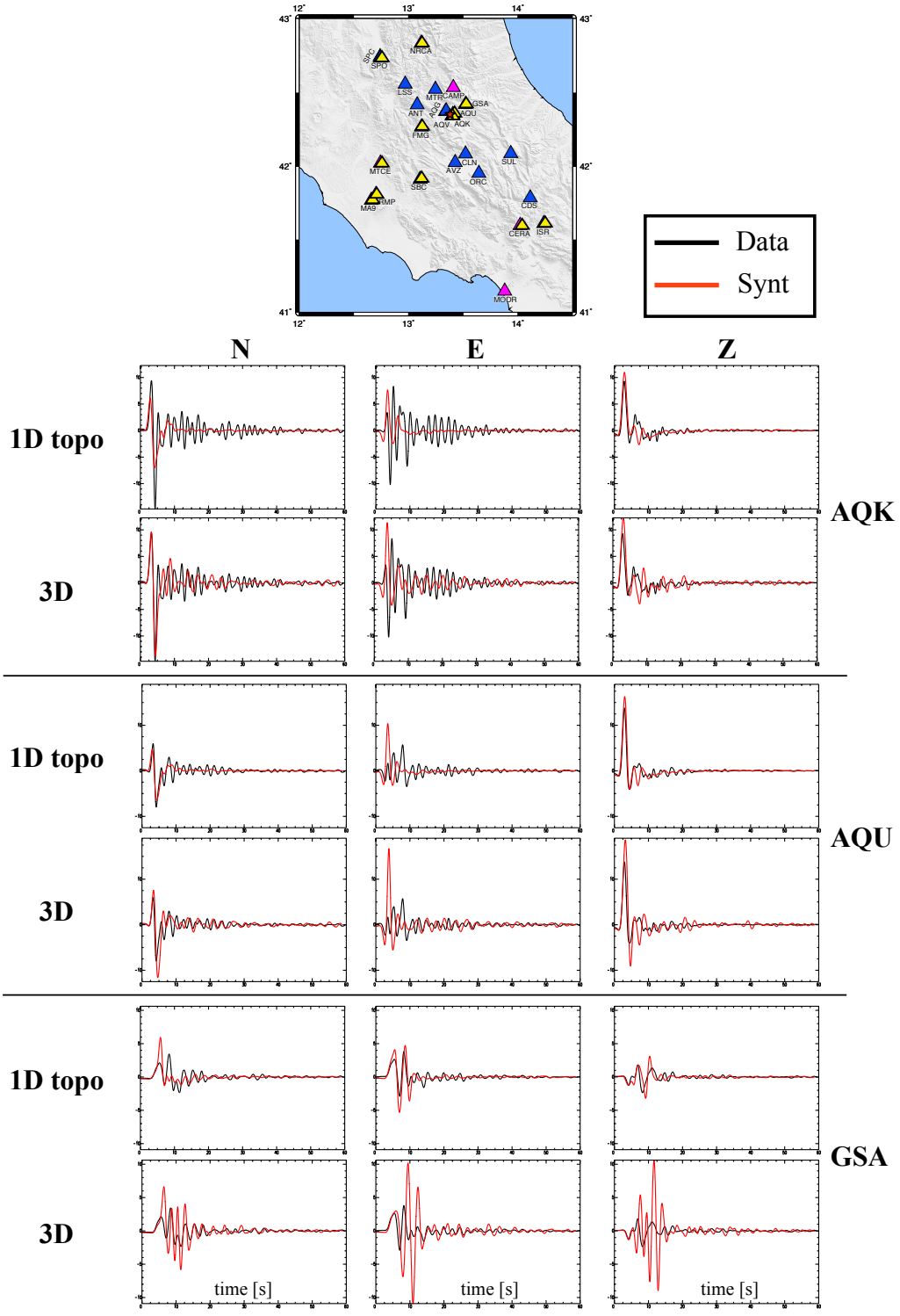


Figure 3.11: Comparison between 1D and 3D models - Three-component velocity (in cm/s) for the indicated stations (yellow triangles in the map). The red star in the map is the L'Aquila event. The data are black, the SEM synthetics are red and refer to the 1D model with topography or to the 3D model (1^{st} and 2^{nd} row for each station, respectively). The filter is 0.02-0.5 Hz and the y axis is the same for each station.

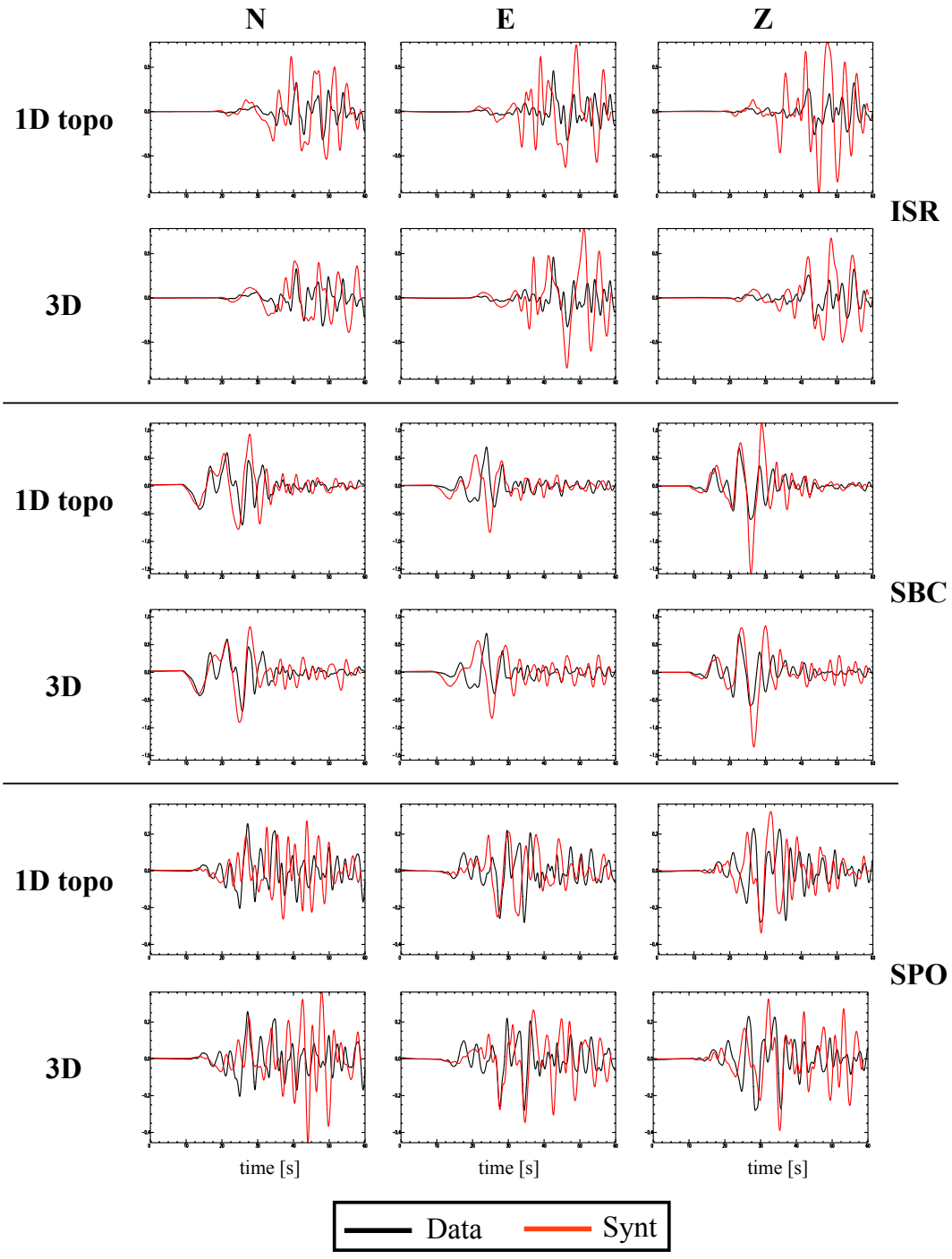


Figure 3.12: Comparison between 1D and 3D models - Same as Figure 3.11.

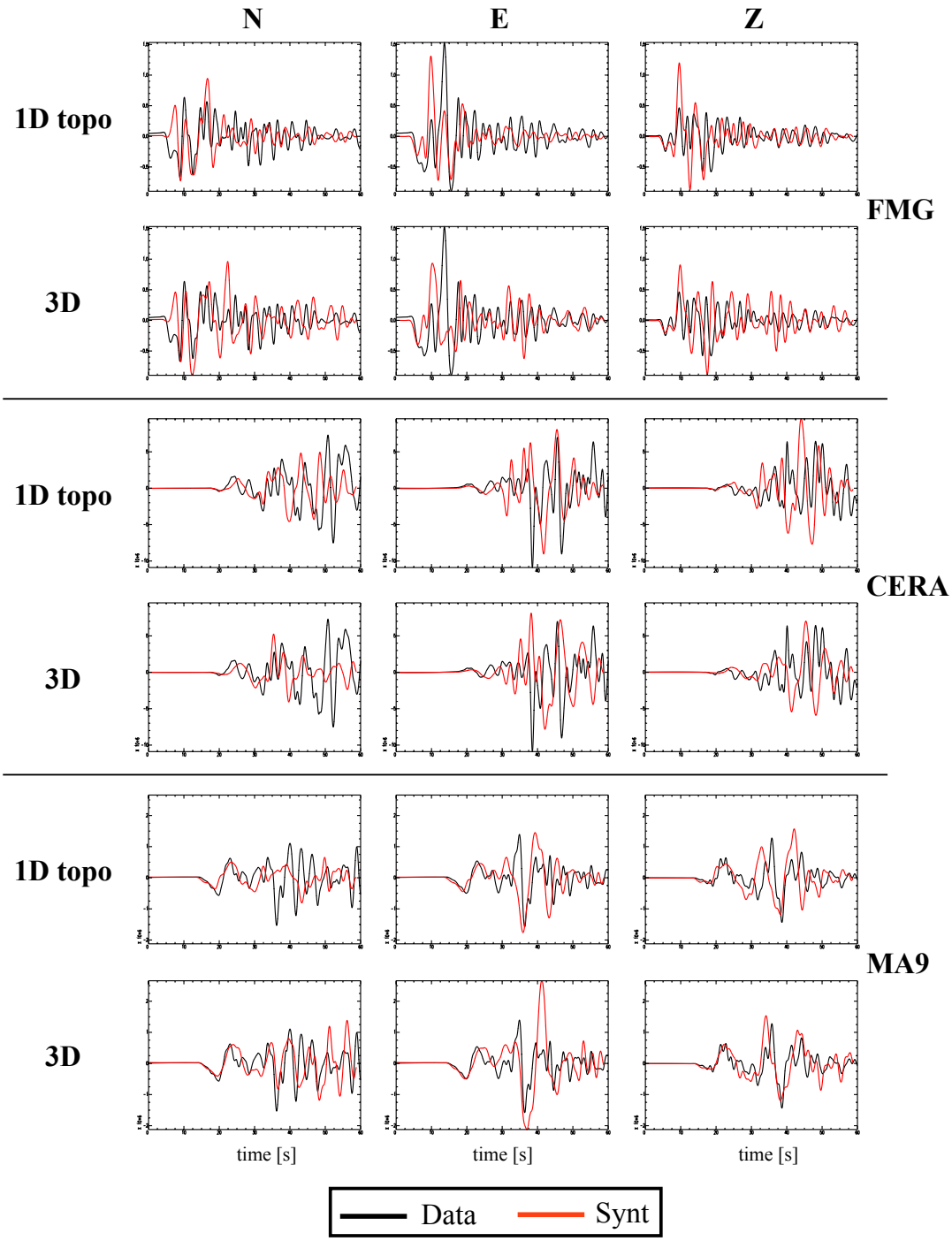


Figure 3.13: Comparison between 1D and 3D models - Same as Figure 3.11. The velocity for FMG is in cm/s, for the other stations is in count/s.

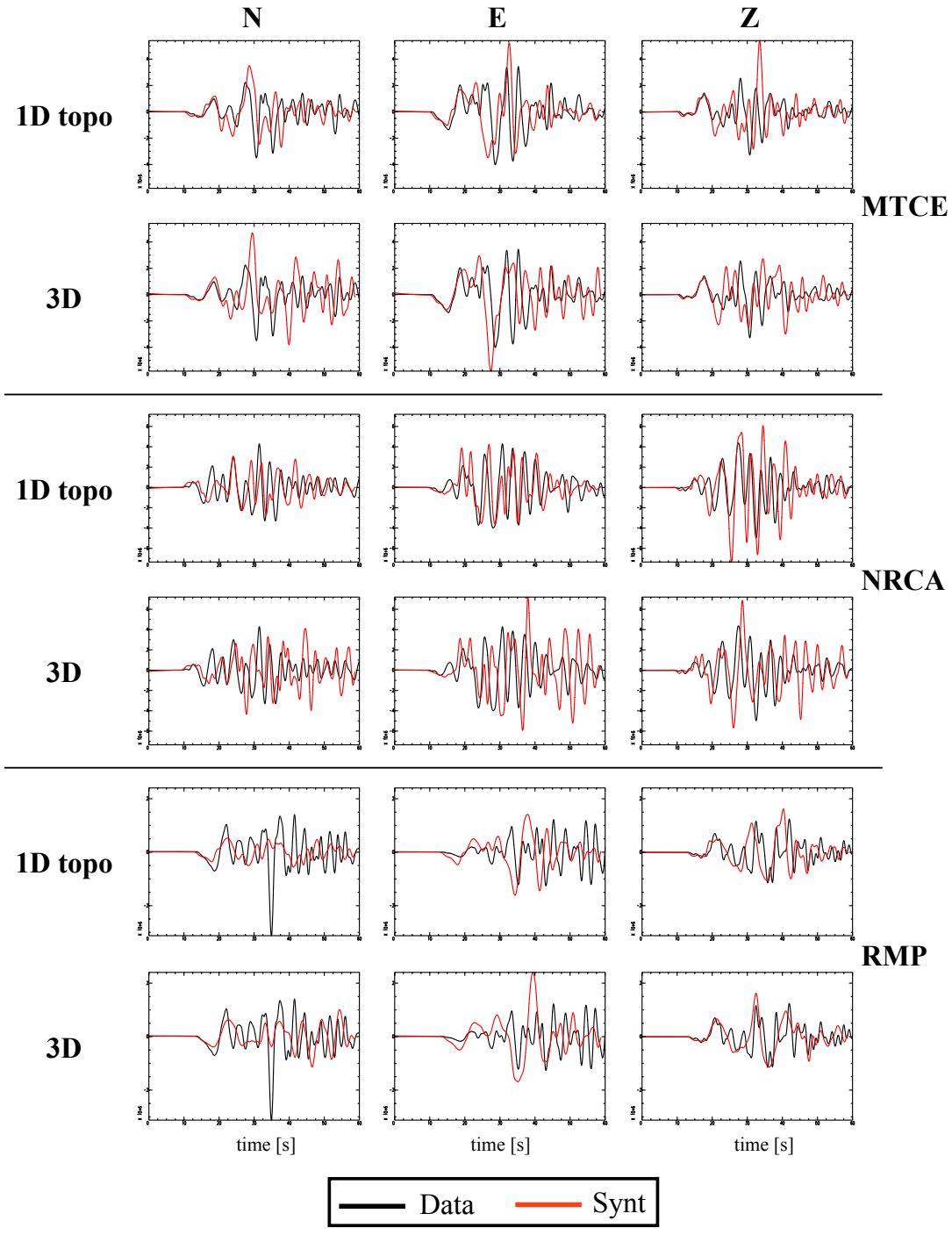


Figure 3.14: Comparison between 1D and 3D models - Same as Figure 3.11, except for the fact that the velocities are in count/s.

to support the conclusion that the first part of the seismograms is mostly influenced by source effects, since only slight changes occur in this part for different structure models.

As shown above, introducing the 3D model, that accounts also for low wave speed sediments, often tends to overestimate the observed data, especially when a simpler 1D model with topography already provides a discrete fit. This may be explained by suggesting that the moment tensor resulting for our model of the finite source is overestimated, producing the strongly amplified picks in the simulated timeseries.

In general, it is worth remarking that, as previously introduced, the results of our simulations, when using the 3D model, are intrinsically affected by a possible incompatibility between source and structure models. In fact, in this case the velocity model is constructed using a 3D tomography. On the other hand, the source model derives from that of Cirella et al. (42) (2009), obtained by an inversion based on a 1D flat velocity profile. This may represent an important source of discrepancy between data and synthetics, that we are not able to solve at present (see section 3.9).

The fact that many features in the observed seismograms are still not or poorly fitted by the adopted 3D model also highlights that a more accurate description of the structure is required. In constructing our 3D model for central Italy we seek to introduce all the available information. However, the adopted velocity profiles are possibly affected by uncertainties, also implicit in the tomographic technique. In addition, many areas are still poorly constrained, as well as the density and attenuation patterns. In particular, the basin effects are accounted for by superimposing to Chiarabba et al. (39) (2010)'s structure the v_s30 layer, that should mimic the low wave speed behaviour of the sediments. However, a refinements of the model in correspondence of these basins is necessary. This would require to define their shape and velocity contrasts at higher resolution, to constrain their depths, and possibly to honor these features with the mesh (e.g., Lee et al. (113), 2008). Nevertheless, presently, this would encounter the limit of the resolvable frequency in our work, as we will show in section 3.7.1. Further analyses of L'Aquila event wavefield may include a new tomography, based upon body wave arrivals, under construction at present (Chiarabba, pers. com.), or an improved 3D model, obtained from an adjoint tomographic inversion as shown in chapter 4.

To conclude, Figure 3.15 shows clearly how the wave propagation in central Italy is strongly influenced by the 3D heterogeneous structure of the region. The wavefront

distortions are due to topographic effects and to the presence of low wavespeed sedimentary basins (drawn in yellow). The ground shaking is prolongated and complicated in and around the basins, where the energy is trapped because of the presence of low velocity sediments. In particular, mainly the short wavelengths are captured by the basins, suggesting that their effect becomes more evident as the frequency increases (section 3.7.1).

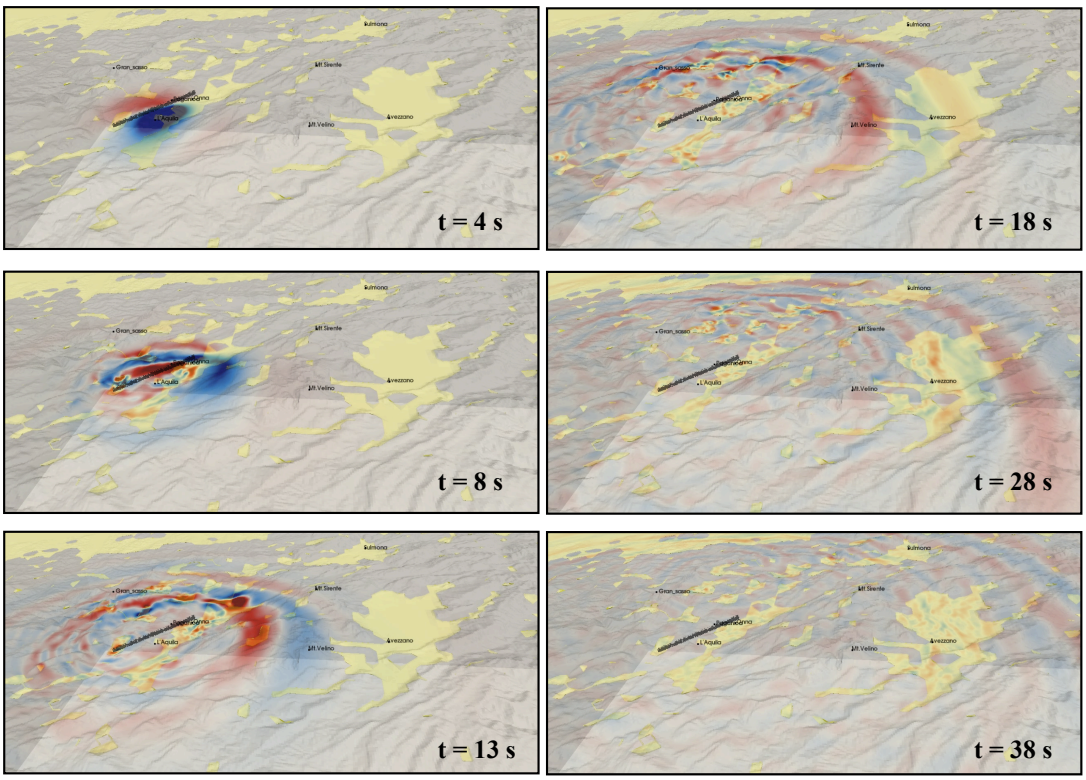


Figure 3.15: Wavefield of the 2009 L'Aquila event - Snapshots at different timesteps of the Z-component velocity wavefield propagating along the surface. Red colors denote positive values and blue negative values. The view is from NW and the black line indicates the fault trace. The wavefront distortions are due to topographic effects and to the presence of low wavespeed sedimentary basins, depicted in yellow, that trap the energy and prolongate the ground motion.

3.7.1 Frequency range analysis

The mesh constructed to implement the 3D velocity model of central Italy allows a resolution up to about 5 Hz (section 3.4.2), and it can be handled by SEM, that guaranties very accurate simulations. Thus, technically we would be able to reach such high frequencies. However, the source model for L'Aquila event derives from that of Cirella et al. (42) (2009), thus, based on how this finite source has been constructed, it effectively represents a low pass filter that limits the frequency working range. Moreover, the considered 3D structure model mainly relies on Chiarabba et al. (39) (2010)'s tomography, that results, as well, in a limited resolvable frequency. Thus, one expects that the final resolution of our modeling is primarily limited by the knowledge, then accuracy, of the adopted source and structure models.

This issue can be studied in more detail by analysing the effect of a frequency range extension up to 1 Hz on seismic waveforms for the L'Aquila event. Figure 3.16 shows, for one of the stations, the comparison between the synthetics for the 3D velocity model (red) and the data (black) for five frequency intervals with an increasing maximum limit up to 1 Hz. The fourth column represents the range 0.02-0.5 Hz used in this work and also by Cirella et al. (42) (2009).

One firstly notes that for frequencies up to 0.17 Hz (i.e., periods down to $\sim$6 s) the synthetics fit the data fairly well. Extending the frequency range, as expected, an increasing complexity delineates mainly in the coda of seismograms (both observed and synthetics). In particular, in our frequency range (0.02-0.5 Hz) multiple late arrivals start to be evident mostly on the horizontal components, and the corresponding 3D synthetics tend to reproduce these features. This means that effects probably related to the presence of low wave speed sedimentary basins begin to influence the waveforms already in this frequency range. Moreover, the considered structure and source models are still able to discretely reproduce such a level of accuracy. When the frequency range extends up to 1 Hz, one immediately notes strong features in the observed seismograms, that the synthetics are no longer able to model. Thus, amplifications and complex resonances, mainly due to basin effects, are more and more enhanced for increasing frequencies, but our simulated wavefields can incorporate only part of these evidences at such high frequency.

The above results are, in general, confirmed by all the stations in the dataset. This leads us to conclude that the structure and source models, presently available to study L'Aquila event, are not yet able to accurately model seismic wave propagation for frequencies higher than ∼0.5 Hz. Thus, as expected, this frequency can be considered as the maximum accuracy reached, at present, by our study.

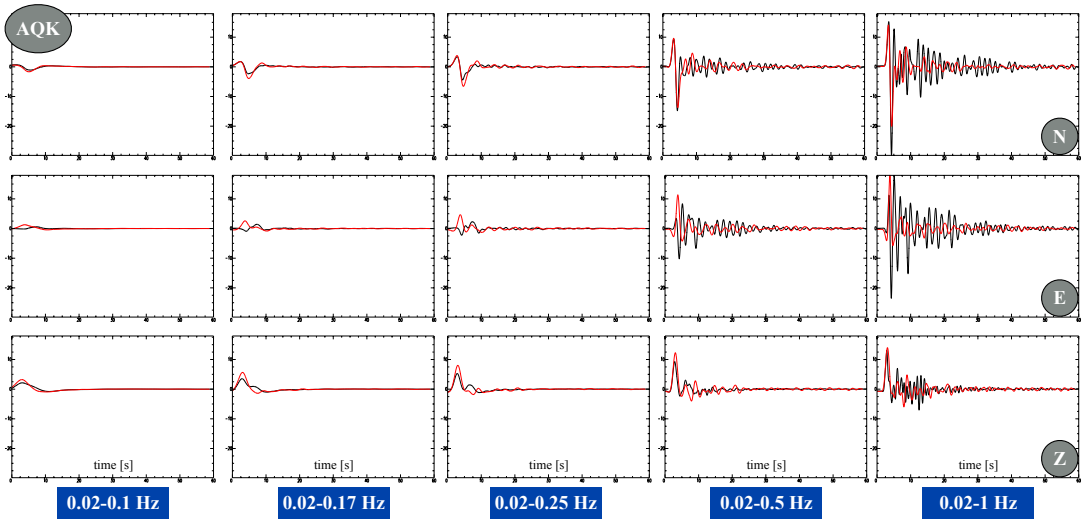


Figure 3.16: 3D model comparisons for increasing frequency - Three-component velocity (in cm/s) for station AQK (see map in Fig. 3.11) and for five frequency ranges (indicated in the blue boxes at the bottom). The data are black, the SEM synthetics for the 3D velocity model are red. The y axis is the same for all seismograms.

To conclude, one can observe synthetics for the 1D model with topography compared to the data in the same frequency ranges. Figure 3.17 shows the same station of Figure 3.16. This example, and the whole dataset alike, show that for relatively low frequencies (up to about 0.17-0.25 Hz) the synthetics behave, with respect to data, as the synthetics for the 3D model. Thus, the three-dimensionality of the structure seems not to significantly affect the waveforms. Increasing the frequencies, instead, the complexity increases and introducing a 3D model becomes essential to capture many features in the observed seismogram coda, otherwise completely missed. This confirms that, in the frequency range considered in our work, a 3D description of the structure is fundamental to accurately reproduce the 3D heterogeneities of the region.

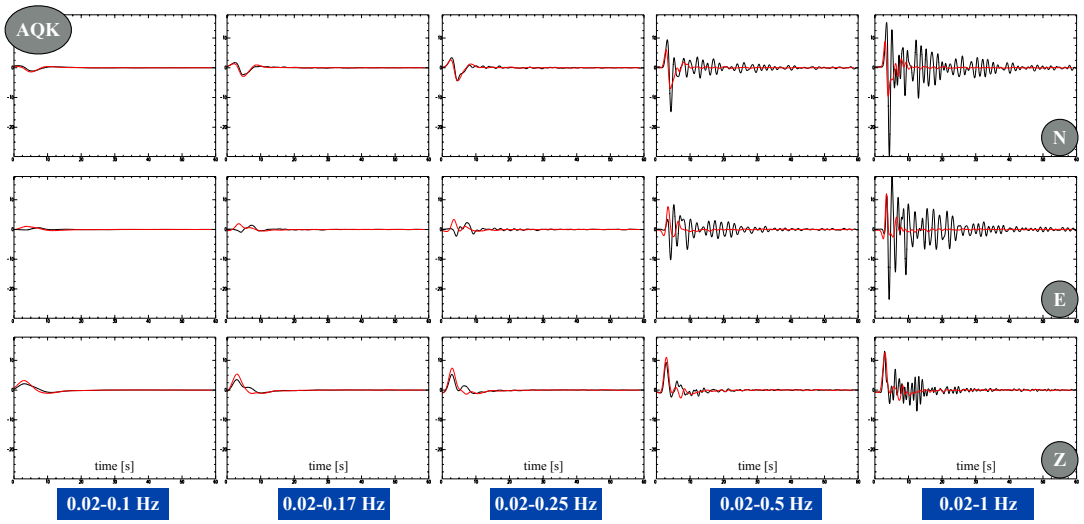


Figure 3.17: 1D model comparisons for increasing frequency - Same as Figure 3.16, except for the fact that the SEM synthetics (red) are obtained using the 1D velocity model with topography.

3.8 Peak ground velocities

Using the code SPECFEM3D, it is also possible to obtain synthetic maps of the peak ground velocity (PGV). These maps can be very useful to reconstruct a hazard scenario for the considered earthquake, highlighting the ground motion behaviour in the surrounding areas.

Many studies evidenced increments of peak ground acceleration and velocity (PGA and PGV) due to the presence of sedimentary basins and also to the source radiation pattern (e.g., Komatitsch et al. (109), 2004; Lee et al. (113, 114, 115), 2008, 2009a,b; Stupazzini et al. (184), 2009, and references therein). Moreover, some of them (Lee et al. (113, 114, 115), 2008, 2009a,b) inferred an increase of PGA and PGV related to the topography (although lower), and this can also contribute to the effects caused by the low wave-speed sediments. Thus, they suggest that also a high-resolution, realistic topographic description should be considered in seismic hazard analysis, especially for densely populated mountainous areas, as advocated by many other studies (e.g., Komatitsch & Vilotte (104), 1998, Massa et al. (137), 2010, and references in both of them). Concerning central Italy, the zone struck by L'Aquila event is classified, according to the normative for the Italian territory, as characterized by high level of seismic

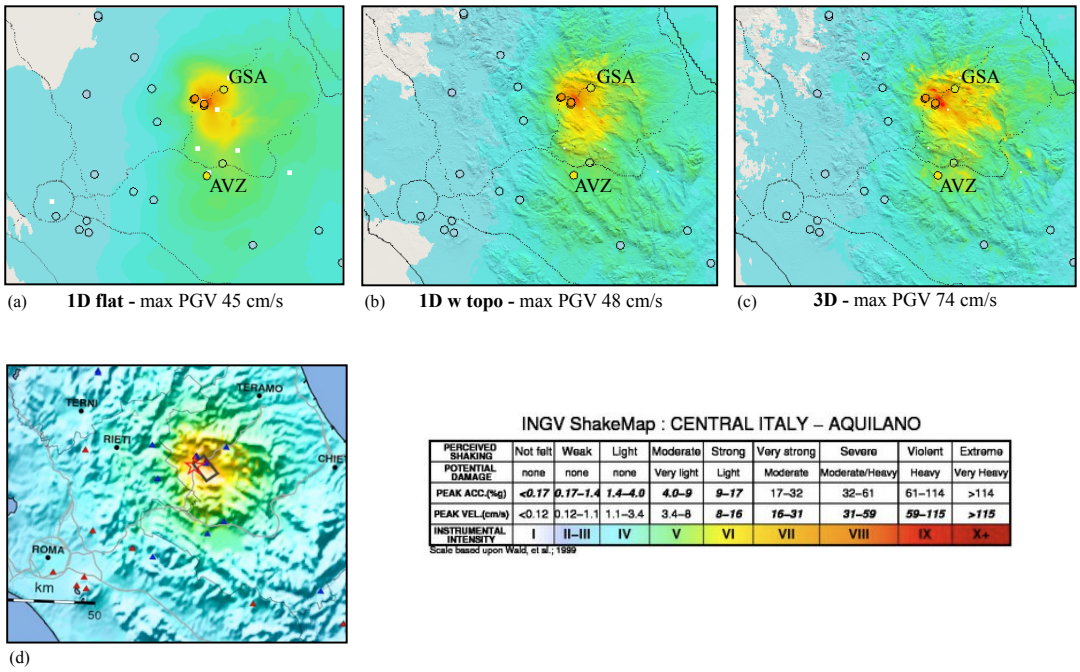
hazard (e.g., MPS (146) Working Group 2004; Akinci et al. (4), 2009). The maximum observed value of PGV in the epicentral area, recorded by strong motion stations of the RAN, is ∼65 cm/s [1]. Moreover, e.g. Akinci et al. (5) (2010) (and references therein) highlighted for L'Aquila earthquake a prominent source directivity effect towards SE, that causes a systematic decrease of PGA and PGV at sites located NW, i.e., in the backward direction of the rupture propagation (compared to sites located SE). This seems consistent with the asymmetric rupture velocity inferred by Cirella et al. (42) (2009). In addition, Akinci et al. (5) (2010) observed higher peak ground velocity for sites located on alluvial sediments with respect to PGV on rock sites, as generally observed in central Italy by previous studies (e.g., Bindi et al. (15), 2004).

This section presents the synthetic maps of PGV, that we obtained for L'Aquila event by considering both the 1D velocity models (with and without topography) and the 3D model, together with the corresponding finite source. Figures 3.18 (a)-(c) show the three maps, allowing us to highlight the effects of introducing the topography and the 3D lateral heterogeneities in the models. The observed values of PGV at the RAN stations are indicated by the circles in the figures, based upon the color scale.

Considering the 1D model without topography (Fig. 3.18 (a)), one notes that in general the observed values of PGV are underestimated in this case. In particular, in the epicentral area the maximum value is 45 cm/s, about 30% lower than the observed one. Introducing the topography (and the attenuation) in the model results in a better agreement between observed and estimated PGV values (Fig. 3.18 (b)). In particular, the topographic ridges are highlighted, e.g. along the Gran Sasso massif in the NE, featuring higher values of PGV with respect to the flat model. An interesting case is station GSA, that is located in a small sedimentary plateau and its PGV value was underestimate by the flat 1D model. On the contrary, now, accounting for differences in the topographic relief, the PGV at this station is better reproduced. The maximum epicentral value is 48 cm/s. Finally, using the 3D structure model (Fig. 3.18 (c)), we obtain PGV estimates in discrete agreement with the values at the stations, and in the epicentral area the maximum PGV is 74 cm/s. In this case the topographic features are still evidenced, and, in addition, the highest values of PGV are mainly observed in correspondence of the sedimentary basins in the region, where most of the energy is trapped. Also the estimate of PGV at station AVZ in the Fucino basin is improved

[1] http://shakemap.rm.ingv.it/shake/2206496920/intensity.html

with respect to the 1D cases above. In all the pictures the source directivity effect on the PGV distribution is evident, with higher values towards SE, however the 3D model is particularly able to reproduce this aspect.



(a) **1D flat -** max PGV 45 cm/s (b) **1D w topo -** max PGV 48 cm/s (c) **3D -** max PGV 74 cm/s

(d)

Figure 3.18: Peak ground velocity maps - (a)-(c) Synthetic PGV maps obtained using SPECFEM3D with the finite fault for L'Aquila event and 3 different velocity models (section 3.4.1): (a) the 1D velocity model without topography, (b) the 1D model with topography, and (c) the 3D model. The circles show the PGV values at the RAN accelerometers. The maximum PGV indicated in each picture is the highest value in the epicentral area. The colors refer to PGV values based on the scale reported on the bottom right. (d) ShakeMap of the ground motion in central Italy generated by the 2009 L'Aquila event (see Faenza et al. (71) (2011) and footnote 1 pag. 99). Considering the scale on the right, the colors correspond to PGV values and also to other parameter ranges, e.g. PGA and instrumental intensity, based upon Wald et al. (204) (1999) (see also footnote 1 pag. 101).

Based upon the obtained results, as expected, the pattern of PGV values results strongly influenced by the inclusion of topography and 3D features in the models. A better agreement with the observed estimates can be obtained only by accounting for all the complexities in the structure. This leads us to conclude once again that an as accurate as possible description of topography and velocity heterogeneities, such as

basins, is essential to model the ground motion behaviour in details.

The synthetic maps produced by SPECFEM3D could be used to construct synthetic ShakeMaps, that would complement the empirically derived ShakeMaps for the ground motion intensity in central Italy (Michelini et al. (140), 2008). For L'Aquila event the empirically obtained ShakeMap (Faenza et al. (71) (2011) and footnote 1 pag. 99) is shown in Figure 3.18 (d). Considering the color scale on the right of this picture, one can observe a remarkable agreement between the PGV pattern and values in Figure 3.18 (d) and those simulated using the 3D model (Fig. 3.18 (c)), especially around the epicentral area.[1]

3.9 Discussion and conclusions

In this chapter we studied the wavefield generated in central Italy by the 2009 L'Aquila event, using SEM, a finite source representation and both 1D and 3D velocity models.

The results of our analyses show that a detailed topographic description and the known 3D heterogeneities of the considered region need to be taken into account. Including all these features in the 3D model, together with a realistic finite fault, we are able to reproduce central Italy ground motion in moderate to good agreement with the observed data. All the recorded waveform components are generally well modeled up to a frequency of 0.5 Hz (i.e., down to a period of 2 s) (e.g., Figs. 3.11 - 3.14).

Remaining discrepancies with respect to data, however, are evident also considering the 3D structure description and the corresponding finite source model. Exploiting the power of present day computational resources and of a numerical technique such as SEM (together with the constructed complex mesh), it would be feasible to very accurately simulate the seismic ground motion, also at very high frequencies. Thus, the principal limitations are caused by the available source and structure models, which knowledge is yet affected by uncertainties and poor constraints. This can bias the synthetic signals and decreases the problem accuracy.

In particular, we have found that the waveforms simulated using the 3D model tend to overestimate the observed seismograms. One can firstly attribute this effect to heterogeneities in the structure not properly accounted for by the adopted model.

[1] Strictly, the intensities less than degree VI of the ShakeMaps are determined from PGA values and therefore the comparison above is not entirely interconsistent. The differences, however, between the intensities determined from PGA and PGV are minor in most instances.

Our analyses show, on recorded wavefields, prominent effects of multiple reflections and amplifications due to topography, as well as to the resonance of low wavespeed sedimentary basins. Comparisons with the synthetics reveal that the description of these basins especially requires further accurate refinements. In fact, at present, their shape and depth are not explicitly implemented in our 3D model, and the low velocities that they feature are taken into account only by incorporating a v_s30 technical layer, that was originally determined for other purposes (Michelini et al. (140), 2008). More generally, a tomographic model at higher resolution seems to be required. To this end, in future analyses it would be useful to test a new 3D velocity model, again based upon body wave arrival times, that is currently under construction (Chiarabba, pers. com.). Moreover, starting from the used 3D structure, improved models may be obtained based upon an adjoint tomographic inversion, as discussed in chapter 4.

One may suggest that another factor that contributes to the amplification of the 3D model synthetics is an overestimate of the resulting seismic moment attributed to the L'Aquila event source. For each of the point sources, that compose the finite fault, we estimate M_0 based on the velocity model, that gives μ, and on the slip, derived from Cirella et al. (42) (2009)'s source inversion. Thus, although the on-fault slip distribution on our source model remains the same as in Cirella et al. (42) (2009), the M_0 distribution changes, influenced by the 3D structure now considered. As a consequence, the amplified phases in the corresponding synthetics can be explained in two ways, likely complementary. A first cause of this effect may be again the adopted velocity model, featuring too high velocities, that produce unreliable high values of μ and then of M_0. On the other hand, one can suggest that the primary cause is an overestimate of the slip. If one considers that adopting a 3D model provides a more reliable description of the structure, instead of a simple 1D profile, thus, lower values of the slip are needed, that, combined with a 3D structure, does not overestimate the data. It is worth noting, moreover, that the amplifications in the synthetics for the 3D model are mainly evident when the 1D model with topography provides discrete fit. Thus, the adopted source model seems more consistent with the 1D profile, whereas when all the 3D heterogeneities are considered this may produce spurious effects. This is quite plausible, since the source inversion is based upon a 1D velocity model. Thus, this discussion clearly highlights that our results in the 3D case are intrinsically affected by an incompatibility between the adopted source and structure models. Calculating

μ from the 3D velocity model seems not to help or even to worsen the effect. This is because the slip is always inferred using a 1D flat model, and, at least in this thesis, the incompatibility is unavoidable, since until now finite source inversions based upon 3D models were not available. Nevertheless, it is worth noting that in future developments of this work, we could test new fault models for L'Aquila earthquake proposed by very recent studies (Trasatti et al. (196), 2011; Volpe et al. (203), 2012) and based on heterogeneous 3D structure models.

As a consequence of the above discussed issues, a tomographic inversion to improve an initial 3D velocity model, as well as a source inversion related to a 3D structure, have strongly attracted our attention and motivated the study presented in chapter 4 and future developments (Appendix A).

An accurate, independent description of density and attenuation is still not possible at present, due to the lack of a comprehensive model of these parameters for the considered central Italy region. Thus, in this work their values are estimated based upon the velocity model and this contributes to increase the uncertainties on the structure definition. Future implementation of more detailed 3D ρ and Q models will allow a more thorough delineation of the characteristics of the structure and a more accurate representation of the ground motion. Especially a proper attenuation implementation is crucial when low wavespeed sedimentary basins are present and site effects are prominent, as in central Italy (e.g., Castro et al. (25), 2004; Bindi et al. (16), 2009).

An important conclusion of our work is that using the 3D structure and the finite fault allows us to model the recorded wavefield also for stations not included in Cirella et al. (42) (2009)'s source inversion. Moreover, despite the required refinements, we are able to capture also features, such as topographic and basin effects, not taken into account by Cirella et al. (42) (2009). These effects, already visible within the frequency range 0.02-0.5 Hz, will become more and more evident increasing the resolved frequency. In fact, most of the basins in central Italy are shallow and relatively small, thus mainly the short wavelengths are trapped. In addition, small-scale topographic features at high frequencies strongly contribute to increase the ground shaking, as seen also in other studies (e.g., Lee et al. (114), 2009a). However, it is fundamental to remark that the structure and source models, presently available and used in this study, work accurately only up to $\sim$0.5 Hz. Thus, the accuracy of the problem is yet limited to this frequency.

This suggests once again the necessity for a structure and source description at higher resolution, if one aims to simulate seismic waveforms in more details.

The analyses of L'Aquila wavefields using 1D velocity models, besides evidencing the need for a 3D model, also enable us a comparison with Cirella et al. (42) (2009)'s results and clearly highlight the effect of topography. In particular, using a finite fault model and a velocity profile similar to those in Cirella et al. (42) (2009), we obtain synthetics in agreement with the seismograms simulated by Cirella et al. (42) (2009), except for few discrepancies probably due to a different source time function. Thus, we conclude that we are generally able to properly implement their inverted source model in our code. Moreover, it results that an explicit incorporation of the surface topography is absolutely required, although the topographic effect could have been partly mapped into Cirella et al. (42) (2009)'s source representation. In fact, some features in the seismogram coda are likely related to topography and they can be captured only by including this surface heterogeneity in the model (section 3.6.2).

As discussed, the effect of topography and lateral velocity variations seems to mainly affect later arrivals in the seismograms. On the contrary, the first part of the timeseries is only slightly influenced by these features and it should be mainly related to the source description. In fact, in this initial portion the seismograms do not significantly change using different structure models (1D or 3D and with or without topography), and they mimic Cirella et al. (42) (2009)'s synthetics. Hence, the characteristics of the constructed source model are crucial to reproduce the wavefield, mainly for the first few pulses. The inverted source of Cirella et al. (42) (2009) is adopted in our work particularly because it has been constrained also using seismic strong motion data. This model provides a quite good fit of the observed seismograms, at least in their first part. However, also there, discrepancies are evident. Thus, in the future, it would be interesting to test other source models, such as the recently updated version of Cirella et al. (42) (2009)'s model (Cirella et al. (43), 2011, submitted) or the model of Scognamiglio et al. (174) (2010). Another option would be to performe a source inversion based upon a 3D model, as already mentioned. For point sources this is presently feasible, as shown in Appendix A. Moreover, using the adjoint method (e.g., Kim et al. (97), 2011) provides valuable results, that also promise for kinematic rupture inversions in 3D structures. In addition, for the L'Aquila earthquake, as cited above, the very recent studies of Trasatti et al. (196) (2011) and Volpe et al. (203) (2012)

propose on-fault slip distributions obtained from inversions based on finite-element methods and 3D velocity models.

Valuable results have been obtained concerning the calculation of the peak ground velocity. In particular, PGV estimates, compared to observed data, dramatically improve considering the 3D velocity model, instead of the 1D models. To this regard, both topographic features and low velocity sediments, causing ground motion amplifications, are highlighted. In addition, characteristics of the source mechanism, such as directivity, become more evident. This confirms that a 3D model, also including topography, contains fundamental information and it is more suited to capture all the heterogeneities of a real Earth structure. Moreover, this suggests that we are able to retrieve important features of the ground motion behaviour, that could be very useful in seismic hazard analyses.

In conclusion, it is worth remarking that our study represents one of the first examples in which the complex wavefield produced by a finite source description in a realistic regional model is analysed and compared to the data. Our work evidences the need for a 3D model that includes also topography, and a finite fault model seems to provide an accurate representation of the source in agreement with the data. Moreover, we prove once again the power of SEM in simulating complex realistic wavefields. The obtained promising results offer the opportunity of future developments, and they can be also considered as preliminary for the adjoint tomographic inversion presented in chapter 4. Finally, we believe that, when improved source and structure models will be available, it will become possible to very accurately reproduce the seismic ground motion, and, as a final result, to obtain precious indications for improved hazard assessment.

Chapter 4

Adjoint tomography of central Italy lithosphere

4.1 Introduction

In this chapter I will present in details the steps and the results of the iterative tomographic inversion, based upon the adjoint method and the steepest descent algorithm, performed to improve an initial 3D model for central Italy lithosphere. The theoretical formulation of the problem is mostly reported in section 2.3, here only the fundamental equations will be highlighted.

To perform wavefield simulations, both forward and adjoint, and to calculate the kernels required in the inversion (sections 2.3.3 - 2.3.4) we used the spectral-element code SPECFEM3D 2.0 presented in section 2.4, that guarantees a highly accurate forward and reverse modeling also for very complex 3D heterogenous models. The accuracy of the numerical method is one of the factors that influences the accuracy of a tomographic inversion, thus using SEM represents a great advantage. In addition, in our inversion we can benefit from all the advantages of the adjoint method (section 2.3) and from its powerful combination with SEM (e.g., Tromp et al. (198), 2008).

The quality of data and of structural and source models, moreover, affects the fit between recorded and synthetic timeseries, and it is crucial in retrieving a physically consistent improved model.

The Italian region is characterized by a moderate to high seismicity (e.g., Scognamiglio et al. (173), 2009) and a large number of broadband stations provides us very

high-quality data. In particular, the 2009 L'Aquila mainshock, analysed in the previous chapter, counts a large number of foreshocks and aftershocks, that have been recorded by both permanent and temporary stations. Thus, we have at our disposal a huge amount of valuable data useful to perform the inversion (see section 4.4).

Concerning the lithospheric structure, the central Italy region features a complex tectonic setting (see section 3.2), evidencing an ongoing NE-trending extension in the Apenninic sector (e.g., Mariucci et al. (130) 1999). In addition, there are several topographic reliefs with a remarkable vertical scalelength, that result in relevant effects on waveform propagation (see section 3.6), and numerous alluvial basins featuring very low velocities of seismic waves that leave strong signatures on the recorded seismograms (see section 3.7). For this region, both 1D and 3D velocity models are available (see section 3.4.1). In principle, a tomographic inversion could start from a well benchmarked 1D model (in our case, e.g., the one used by Cirella et al. (42) (2009)), that would reduce the computational costs associated to the inversion, also allowing for the calculation of both the gradient and the Hessian of the misfit function to be minimized (section 2.3.4). However, performing a tomographic inversion based upon initial 3D models and full 3D simulations is becoming feasible (as introduced in section 1.1). Moreover, the study of the L'Aquila event in the previous chapter has highlighted the importance of introducing 3D models, that allow us to account for a huge amount of valuable information contained in the seismograms, neglected otherwise. These are the key points that motivated our choice to start the adjoint tomography of central Italy from a 3D model of the region which includes as much known features as possible (see sections 3.4.1 and 4.2.1).

Regarding the source models, it would be desirable to describe them with parameters derived from an inversion based upon 3D models. At present, for central Italy the available source mechanisms of the events are derived through inversion techniques that use 1D velocity models (section 4.3). Thus, in this tomographic inversion we started by using these solutions. Once an updated 3D model will be obtained, it could be used to improve the source description with a 3D-based inversion (e.g., Liu et al. (119), 2004; Kim et al. (97), 2011; Appendix A).

The choice of the misfit function to be minimized is another crucial factor, since it determines the measure of the difference between data and synthetics we are dealing with. This results in adopting a specific tomographic approach and only the features

to which the considered measurements are sensible will be imaged by the inversion. Several choices are possible (e.g., Tromp et al. (197), 2005) and in this work we choose to minimize (multitaper or cross-correlation) traveltime differences between observed and synthetic waveforms (sections 2.3.3 and 4.6). Thus, mainly the model perturbations affecting the shift of phase arrivals will be accounted for. Instead, only slight adjustments will result in the timeseries amplitude differences. This may suggest the need to perform also an amplitude tomography, and, above all, it reveals the necessity of inverting also for source parameters, whose errors are inevitably mapped into structural models when a tomographic inversion is performed. These aspects will be the objective of future developments of this work.

Finally, the efficiency of the chosen minimization algorithm defines the number of iterations required to approach the minimum of the misfit function, affecting consequently the computational costs associated to the procedure. As mentioned, we choose to use the steepest descent method (in combination with adjoint method) presented in section 2.3.4.1. Future extensions of this work may include a comparison with a conjugate gradient method (e.g., Tape et al. (187), 2007), in order to address to a fast convergence of the minimization procedure.

Next sections 4.2-4.9 are organized as the steps of an iterative tomographic inversion.

- We start by defining the initial velocity model $\mathbf{m}^0$ (section 4.2.1) and seismic source description (section 4.3), together with the observed dataset (section 4.4).

- Then, we discuss the algorithm to select the windows of recorded and synthetic timeseries (section 4.5.2), and the chosen measurements and misfit functions estimated in each time window, that are required for the inversion (section 4.6).

- Next, we present the construction of the gradient of the misfit function, i.e. the misfit kernels (section 4.7.1), that is used to calculate the new model $\mathbf{m}^1$ updating the initial model $\mathbf{m}^0$ (section 4.8).

- Finally, the two previous steps can be repeated to obtain the model at iteration n, $\mathbf{m}^n$, by updating the one at the previous iteration, $\mathbf{m}^{n-1}$, now considered as starting model. Thus, performing successive iterations, the initial velocity model can be improved, while minimizing the misfit function. In sections 4.8 - 4.9 we discuss the results of the inversion procedure, starting from the first update $\mathbf{m}^1$

of the initial velocity model $\mathbf{m}^0$, until the model $\mathbf{m}^5$, that has been obtained by the last iteration performed in this work.

4.2 Initial model

Velocity structure and source models affect seismic wave propagation and, therefore, the errors associated to the respective parameters are at the origin of the discrepancies between observed data and simulated synthetics. Thus, ideally, one should perform a joint source-structure inversion in order to improve simultaneously both source and structural models, reducing their effect on the total misfit (e.g., Pavlis & Booker (157), 1980). An example of this joint minimization procedure, based upon SEM and adjoint method, is in Tape et al. (187) (2007) for a synthetic 2D experiment. For realistic 3D cases, this would require a remarkable computational effort, thus, one can also start by inverting for only one model, structure or source, while keeping the other fixed. Updating only one model, for example the structure, one assumes that the source inaccuracies are not affecting the waveforms, but obviously the errors on the source parameters are implicitly mapped into the structural model. Thus, once the structural model has been firstly improved, it can be used to deal with source parameter inversion, and so on, seeking to minimize the combined effect of source and structure on the difference between data and synthetics.

In this work, as anticipated, we followed the last approach, initially inverting only for the structural model of central Italy, while keeping the source description unchanged. Next section (4.2.1) describes the initial 3D velocity model of the region, that will be updated through the adjoint gradient-based inversion. The SEM mesh constructed to discretize this model is described in section 4.2.2. It is worth noting that the geometry of the mesh remains fixed for all the iterations of the inversion procedure, whereas the material properties associate to its GLL points (section 2.4) change accordingly to the improved model. The representation of earthquake sources used for the whole tomographic inversion is, then, presented in section 4.3.

4.2.1 Initial 3D velocity model $\mathbf{m}^0$

The first step in defining the initial model is to choose the variables for its parametrization. In this adjoint tomography, we consider the following isotropic model parameters:

- Compressional wave-speed v_p,
- Shear wave-speed v_s.

These are the only variables directly updated in the inversion procedure, and, formally, to use the steepest descent equation 2.125 for the model update, they need to be transformed into log-normalized quantities [1] (Tarantola (193), 2005), that is

$$V_p = \ln\left(\frac{v_p}{v_{p_0}}\right), \qquad V_s = \ln\left(\frac{v_s}{v_{s_0}}\right), \tag{4.1}$$

where v_{p_0} and v_{s_0} are scaling parameters with dimension m/s used to non-dimensionalize v_p and v_s, respectively. The two variables V_p and V_s are defined at each GLL point of the discretized volume and represent the components of the model vector $\mathbf{m}$ [2]. Their variations are

$$\delta V_p = \delta\left[\ln\left(\frac{v_p}{v_{p_0}}\right)\right] = \frac{\delta v_p}{v_p}, \qquad \delta V_s = \delta\left[\ln\left(\frac{v_s}{v_{s_0}}\right)\right] = \frac{\delta v_s}{v_s}, \tag{4.2}$$

usually denoted as $\delta \ln v_p$ and $\delta \ln v_s$, respectively (e.g., eq. 2.89).

The mass density ρ is also included in the structural model. Fixed v_p and v_s, this variable has a relatively low sensitivity to the traveltimes we seek to fit, thus, it has not been considered as a model parameter, strictly speaking. However, the definition of ρ is required to practically run the code SPECFEM3D, thus its values are determined from v_p using the relationship (eq. 3.3 section 3.4.1)

$$\rho = 0.025\, v_p^2 - 0.055\, v_p + 2.134, \tag{4.3}$$

where ρ is in g/cm^3 and v_p in km/s. At each iteration, once v_p is updated, the new values of the density for each point are calculated using equation 4.3 and are used in the code to allow for simulations.

Concerning the attenuation, it is implemented as described in sections 3.4.1, 3.4.4. Thus, the quality factor at each GLL point is estimated, based on Olsen et al. (151) (2003)'s model, relating Q and v_s by Q=0.02 v_s (v_s in m/s; eq. 3.1), and it can assume discrete values within 40 and 150, with step of 10. Below the Moho we set Q=9000, i.e., no attenuation. As for the density, the definition of the quality factor is necessary to

[1] This is possible since v_p and v_s are positive quantities.

[2] In general, the model vector is written as $\mathbf{m} = [V_{p\,1}, ..., V_{p\,k}, ..., V_{p\,N_{GLL}}, V_{s\,1}, ..., V_{s\,k}, ..., V_{s\,N_{GLL}}]^T$, with N_{GLL} the number of GLL points in the considered volume (e.g., Tape et al. (189), 2010).

run the simulations if the attenuation is incorporated. Thus, Q is not directly updated, but its new values are calculated at each iteration from the updated v_s.

Anisotropy has not been incorporated in the structural model.

Once the parametrization has been fixed, in order to start from the most complete available description of the structure, we choose to assign the values of the considered parameters v_p and v_s based upon a 3D model of central Italy lithosphere, as introduced in 4.1. We use the laterally heterogeneous tomographic model employed for L'Aquila mainshock simulations and described in section 3.4.1. However, as discussed in paragraph 3.4.4, 3D model simulations for L'Aquila require a huge amount of computational time. Since in the tomographic inversion several events are involved and a large number of simulations needs to be run, we choose to reduce the resolution of the problem by simplifying the corresponding mesh (see section 4.2.2). As a consequence, the minimum wavespeed resolved by numerical simulations increases, and, then, in the initial 3D tomographic model used in the inversion, we limit the minimum v_p and v_s to be, respectively (cfr. eq. 3.2)

$$v_p^{min} = 1753.3 \text{ m/s}, \qquad v_s^{min} = 984.5 \text{ m/s}. \tag{4.4}$$

All the other characteristics of the model are the same as in section 3.4.1. We label the initial velocity model for the inversion as $\mathbf{m}^0$.

Figure 4.1 shows 3D views and sections of the initial models for v_p and v_s. The areas depicted in red feature very low wavespeeds and are related to the alluvial basins, that are present in central Italy and are accounted for by using the v_s30 model (see section 3.4.1).

4.2.2 Mesh

As anticipated in the previous paragraph, it is advisable to reduce the computational costs when dealing with simulations of several events. Moreover, the study of L'Aquila mainshock (chapter 3) revealed that a resolution of 5 Hz, that would be provided by the complex mesh described in section 3.4.2, is considerably larger than the resolution allowed by the considered source and structural models, that finally limit the accuracy of the simulations up to about 0.5-1 Hz (i.e, down to ∼1-2 s). All these aspects, lead us to construct a new simpler mesh for the tomographic inversion in central Italy, with a resolution calibrated upon the estimated accuracy of the problem.

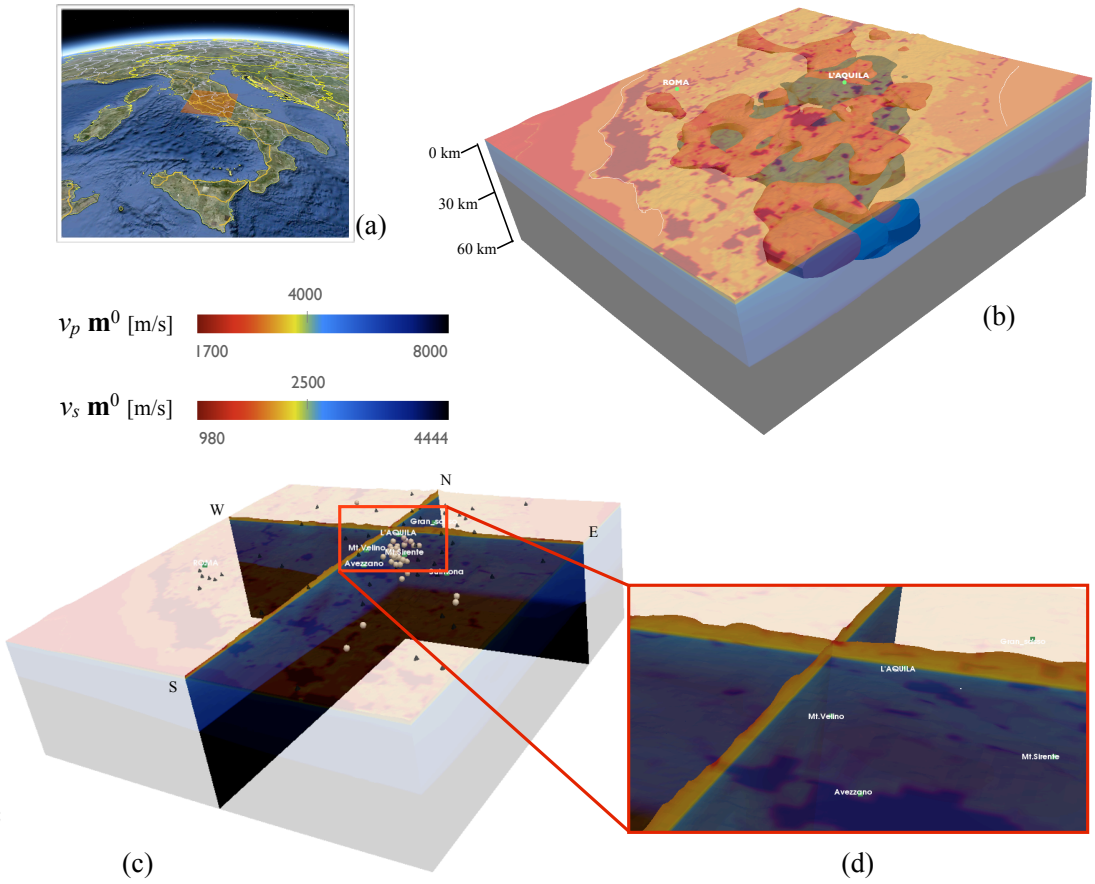


Figure 4.1: Initial 3D velocity model $\mathbf{m}^0$ - (a) View of the simulated area in central Italy. (b) Three-dimensional view from SE of the initial tomographic model $\mathbf{m}^0$ in central Italy. The scales for v_p and v_s are on the left. The considered volume is 200 km x 200 km x 60 km and in black is represented the homogeneous layer below the Moho surface (Di Stefano et al. (60), 2011). In transparency the internal tomographic model is evident. Blue features correspond to v_p=6400 m/s, orange features correspond to v_p/v_s=1.84. (c) W-E and N-S cross-sections of figure (b). Black triangles and white dots denote the receivers and the events considered in the inversion procedure, respectively (see Fig. 4.3). (d) Close-up of the red inset in figure (c). In red are depicted low wave-speed areas, often associated to sedimentary basins present in central Italy.

The mesh built in this case is represented in Figure 4.2. As the mesh for L'Aquila simulations (section 3.4.2), it discretizes a volume of 200 km x 200 km x 60 km in central Italy, with limits of latitude and longitude, respectively, (41.10 - 42.90)°N and (12.04 - 14.45)°E (Fig. 4.3 (a)). However, the size of the hexahedral elements on the top is about 1 km, thus the numerical resolution of the mesh now reaches ∼1 Hz (i.e., a minimum period of 1 s). To account for velocity increase, then, the size of the elements is triplicated along depth. Using this mesh, one forward simulation of 60 s requires about 40 minutes on 256 cores, compared to a forward simulation with the complex mesh for L'Aquila event, that required about 32 hours on 308 cores (section 3.4.4).

The mesh honors the free surface topography, that is imported using Shuttle Radar Topographic Mission (SRTM) data (converted to UTM coordinates) with an original resolution of 90 m (Jarvis et al. (90), 2008), then smoothed at 1 km. The topography of the Moho discontinuity as described by Di Stefano et al. (60) (2011) is also honored and it is represented by the top of the blue layer in Figure 4.2.

The mesh contains 194'400 hexahedral elements with a total number of GLL points equal to 13.5 x 10^6. Initially, to each of these GLL nodes is assigned a value of the parameters v_p and v_s (thus, of the corresponding V_p and V_s in 4.1, and of ρ and Q) based upon the 3D tomographic model considered for central Italy (section 4.2.1). At each iteration, the geometry of the mesh remains the same, but the values of the model parameters at the GLL points will be replaced by the updated ones [1].

4.3 Earthquake sources

In order to perform the tomographic inversion for central Italy, we take advantage of the huge amount of events recorded before and after the 6th April 2009 L'Aquila mainshock. Our model of the region has an extension of 200 km x 200 km x 60 km (see section 4.2.2), thus we started by considering all the events occurred in this volume from January 2009 and September 2010 with local magnitude $M_l \geq 3.0$. Among them, we finally selected 63 events with available seismic source solutions and at least few

[1] Based upon the description of the code SPECFEM3D in section 2.4, this means that at each iteration new updated values of the model parameters will be assigned to the GLL points, whose coordinates are fixed in the mesh databases. Note that, if the model contains any discontinuity, two different values of each model parameter are assigned to the gridpoint that lie on the interface, one for each side of the interface.

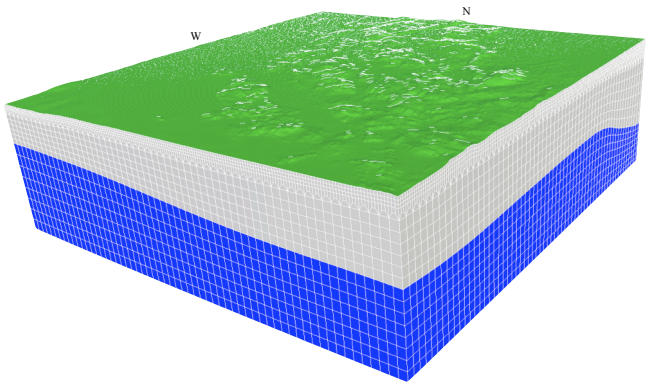


Figure 4.2: Hexahedral mesh for central Italy tomography - The mesh discretizes the volume in Figure 4.1 (b) in central Italy and contains 1.9 x 10^5 hexahedra. It honors the free-surface topography and the Moho surface (Di Stefano et al. (60), 2011), that is the top of the blue layer. The Moho is at about 30 km depth in the NW direction and deepens moving toward SE and NE.

measurements for the shortest considered period range (2-20 s, section 4.5.1). They occurred between March 2009 and August 2010 with $3.2 \leq M_l \leq 5.3$ and are shown in Figure 4.3 (a) based upon their depth. In total, the foreshocks of the L'Aquila event are 3, whereas the aftershocks are 55.

Each event is represented as a point source and, in order to be implemented in the code SPECFEM3D for simulations, it is described by the following Centroid Moment Tensor (CMT) parameters (see section 2.4):

- Coordinates of the hypocentre, i.e., latitude, longitude and depth;
- Six components of the moment tensor;
- Half duration *hdur*, i.e., the half width of the Gaussian that approximates the point source time function (section 2.4);
- Origin time.

The source solutions for the considered events are obtained by using a Time Domain Moment Tensor (TDMT) technique (Dreger & Helmberger (64), 1993; Scognamiglio et al. (173), 2009). This method has been recently implemented at Istituto Nazionale

di Geofisica e Vulcanologia (INGV) and allows rapid determinations of moment tensors from broadband waveform recordings in the period range 10-50 s (Scognamiglio et al. (173), 2009). The high-quality data recorded by the Italian National Seismic Network (INSN) and the Mediterranean Very Broadband Seismographic Network MedNet, together with a laterally homogeneous, layered velocity model, are used to obtain the solutions for events with $M_l \gtrsim 3.0$ [1]. In the TDMT procedure each moment tensor solution is associated to an optimal depth, that is fixed and not explicitly inverted. Latitude and longitude of the events are the revised locations provided by INGV[2].

Another option would have been to use the solutions available in the European - Mediterranean Regional Centroid Moment Tensor (RCMT) catalog (e.g., Pondrelli et al. (162), 2006) for earthquakes with $4.5 \leq M_l \leq 5.5$. The RCMT calculation allows rapid determination of source mechanisms and it is based on an iterative inversion of long period ($\geq$ 35 s) surface waves recorded at regional distance by primarily MedNet stations. The 1D velocity model PREM is used (Dziewonski et al. (66), 1981). However, most of our considered events are too small and only the TDMT solutions are available, thus, for sake of consistency, we choose to use the TDMT mechanisms for all the earthquakes.

Figure 4.3 (c) shows the distribution of moment magnitude M_w for the 63 events obtained by using the TDMT technique. The values are comprised in the range $M_w \in [3.2; 5.4]$, but only four events have $M_w \geq 5.0$. The depth distribution is represented in Figure 4.3 (d), showing that the values range from 2.0 to 18.0 km, but are mainly concentrated between 6.0 and 10.0 km depth.

To simulate a point source using SPECFEM3D one can, in principle, use a Gaussian source time function with a zero width, i.e., a delta function centered at the CMT time. After the simulation, a convolution with an appropriate source time function and/or filtering operations would remove the noise caused by a moment-rate delta function. However, we prefer to assign to each event a non-zero half duration *hdur* calculated from the TDMT scalar seismic moment M_0 by using the relationship (Dahlen & Tromp (51), 1998)

$$hdur = 2.4 \,\mathrm{x}\, 10^{-6}\, M_0^{1/3}\,, \tag{4.5}$$

[1] http://cnt.rm.ingv.it/tdmt.html
[2] available at http://cnt.rm.ingv.it/

where *hdur* is in second and M_0 in dyne x cm. Figure 4.3 (e) summarizes the *hdur* values for the 63 events. Most of the earthquakes have half durations < 0.4 s, which is small compared to the minimum period of 2 s considered in the inversion and compared to the minimum threshold $\sigma_0 = 1$ s that we assign to the measurement uncertainty (eq. 4.20). Under these conditions, the approximation of earthquakes as point sources can be considered reliable.

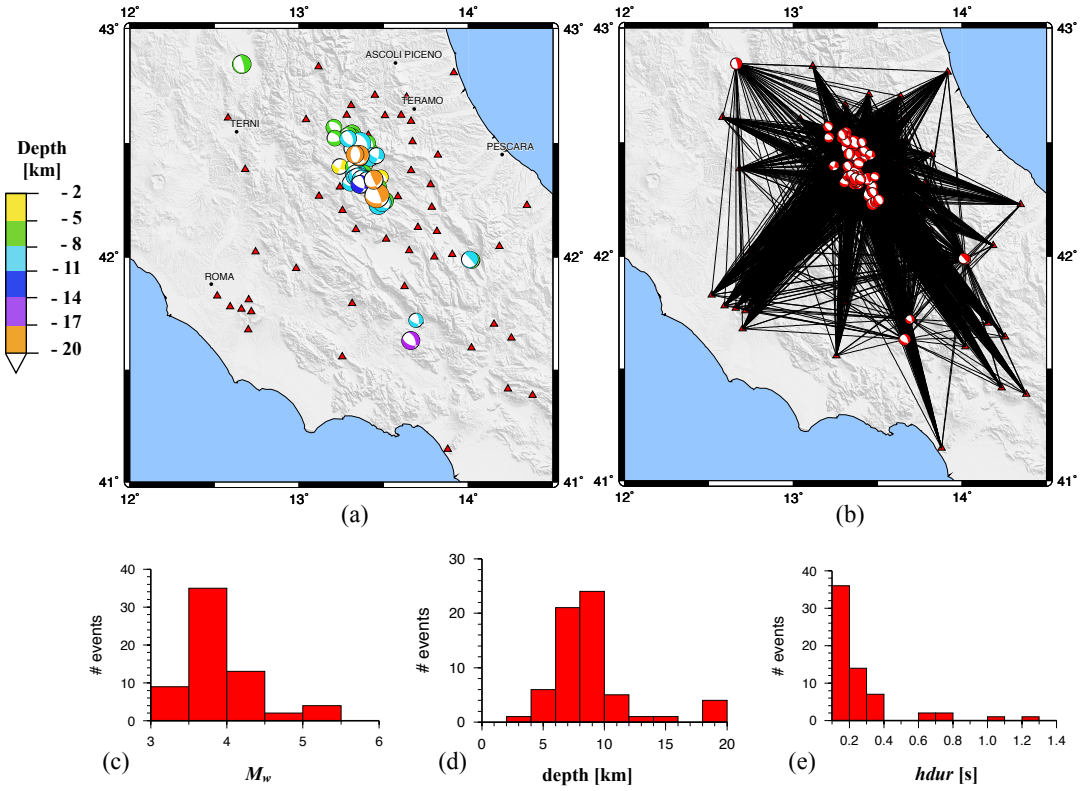


Figure 4.3: Events and receivers for central Italy tomography - (a) The 63 events and the 52 velocimetric stations (red triangles) used for the adjoint inversion. The beachballs have dimensions proportional to event magnitudes and colors referred to the depth color scale on the left. (b) Ray coverage considering all the events and the stations. (c) Moment magnitude M_w distribution for the considered 63 events. The values of M_w are obtained with a TDMT technique (Scognamiglio et al. (173), 2009). (d) Distribution of the depth values for the considered events. (e) Distribution of the half durations *hdur* estimated by using equation 4.5 in the text (Dahlen & Tromp (51), 1998).

As stated above, the source model parameters have been kept fixed for all the performed iterations, improving only the structural parameters. However, we did some

preliminary tests also on source inversion, by using an automated procedure developed by Liu et al. (119) (2004), that allows moment tensor and earthquake location improvements. The technique, summarized in Appendix A, has a remarkable computational cost, thus we won't deal with this issue in this work, postponing it to future developments. This is also justified by noting that the initial misfit between data and synthetics has to be mainly ascribed to perturbations in the structural model parameters. Thus, source inversion will be tackled once an improved tomographic model is obtained after few iterations of the adjoint inversion procedure.

4.4 Observed data

The objective of the inversion procedure is to minimize the difference between the data and the synthetics, quantified by a misfit function. Thus, the seismograms simulated by using the numerical technique have to be compared with the observed seismograms. To this end, we considered as data the seismic waveforms generated by the 63 events presented in section 4.3 and recorded by 52 three-component velocimetric stations. The map of the receivers is shown in Figure 4.3 (a) together with the considered events. Figure 4.3 (b) represents the ray coverage based upon the 63 events and the 52 stations.

In order to select the receivers suitable for the inversion, we started by considering all the stations that recorded each event. Among them, we selected only the receivers inside the simulated volume and we neglected the timeseries that presented some gaps in the recording. The total number of stations finally chosen is 52 and each event has been recorded by at least 21 of them and at the maximum by 44 of them.

Most of the considered 52 receivers are digital permanent seismic stations of the INSN managed by INGV [1]. Moreover, there are several broadband stations belonging to the MedNet network[2]. All these 52 stations provide very high-quality data useful for the inversion. (For more details on the networks see section 3.3).

Almost all the 63 events have been also recorded by some temporary velocimetric stations, especially the aftershocks immediately following the main event of L'Aquila. However, the responses of these instruments refer to very short periods (mainly up to 1

[1] Seismograms and receiver responses available at http://eida.rm.ingv.it/

[2] http://mednet.rm.ingv.it/data.php

s) corresponding to very high frequencies. Thus, in this work we choose not to include these stations that could be useful for successive studies at higher resolution.

Finally, 14 events out of the 63 have been also recorded by accelerometric stations of the national strong motion network (RAN)[1] (as for L'Aquila mainshock, section 3.3). However, the data from these stations have not been considered for central Italy tomographic inversion. In fact, the processing required by accelorometric seismograms is different compared to the velocimetric data processing (see section 3.5). This may influence the corresponding adjoint sources estimated within each time window selected in the seismograms, since their formulation involves processed time series (see, e.g., eq. 2.96 and Table 4.1). Thus, at present, we prefer to consider only the velocimetric seismograms, all processed at the same way, in order to produce adjoint sources for kernel calculation that are consistent to each other.

4.5 Forward modeling

4.5.1 Forward simulations and data processing

The adjoint tomographic inversion we are performing requires to produce, at each iteration, the synthetic seismic wavefields for the given model. They are necessary to estimate the misfit functions by comparing them to the observed data (section 4.6).

As a starting point of the procedure, we have considered as input of the code SPECFEM3D (1) the initial velocity model $\mathbf{m}^0$ (section 4.2.1), (2) the corresponding mesh (section 4.2.2) and (3) the TDMT source descriptions (section 4.3), and we have performed a forward simulation for each of the 63 considered events (Fig. 4.3 (a)).

The synthetic seismograms are obtained for the three components (East, North, vertical) of each of the 52 stations that recorded the events (section 4.4 and Fig. 4.3 (a)-(b)). The simulations are set to produce 60 s long timeseries with a time step of 1 x 10^{-3} s (i.e., the sampling rate of synthetic seismograms is originally 10^3 Hz).

Before comparing data and synthetics to select the measurement windows (section 4.5.2), the time series, both recorded and simulated, need to be processed exactly in the same way. The observed data used for our tomographic inversion are three-component velocity seismograms, that include the instrumental response (section 4.4). We consider

[1]Data available at http://itaca.mi.ingv.it/ItacaNet/

the synthetic seismograms in displacement and perform the processing summarized in what follows:

1. Resample data and synthetics at 100 Hz;
2. Cut data and synthetics between -10 and 59 s;
3. Deconvolve the corresponding instrumental responses from the data, while transferring them into displacement (in m) and filtering them in two period bands, (2-20) s and (6-20) s;
4. Apply the same transfer filter to the synthetics between (2-20) s and (6-20) s, while keeping them in displacement (in m);
5. Rotate East and North components into radial (R) and transverse (T) components.

Note that, as required, all the operations applied to the data have been applied to the synthetics as well, except for the deconvolution of the instrumental responses, that is necessary only for the data. However, the same transfer function used for the data is applied to filter the synthetics instead of a Butterworth filter, thus the filtering operation is consistent in the two time series.

The period ranges considered in the inversion are 2-20 s and 6-20 s. This choice is primarily based on the instrument characteristics and the estimated accuracy of the problem (down to about 1-2 s), due to the considered source and structure models[1]. Moreover, considering multiple filtering bands allows us to highlight, through the kernels (section 4.7.1), the sensitivity of the waveforms to features that arise at different frequencies, and thus to better sample the model. Generally, the two period bands show both body and surface waves, and effects of the alluvial basins or of the topography are often visible (see Figs. 4.26 - 4.37). Thus, many interesting phases are evident in both period ranges and successive iterations will lead the synthetics to reproduce them more and more accurately, while enriching the model with the corresponding features.

As the resolution of the model increases through the inversion procedure, the period range may be extended in order to fit new features. However, for the iterations performed in this work we preferred to keep the period bands fixed.

[1]The accuracy of the simulations, in this case, has been tuned to the problem resolution, by constructing an ad-hoc mesh (section 4.2.2).

4.5.2 Time-window selection

After producing the synthetic seismograms and processing them together with the data, the next step consists in comparing all these timeseries, to select the time windows suitable for measurements. This is required at each iteration of the procedure.

In principle, any time window within which data and synthetics have a significant signal to noise ratio and match reasonably well can be considered for measurements. In this work, we have used the powerful code `FLEXWIN` [1] (Maggi et al. (125), 2009), that allows for an automated selection of the windows based upon these criteria. The algorithm is highly flexible and can be applied to both local, regional and global seismological scenarios. Moreover, it can handle timeseries from 1D simulations, but also seismograms for 3D models, that usually include more complex phases. In fact, one of the great advantages in using this code is that one does not need to identify and label specific phases in the seismograms. The algorithm, only accounting for some user parameters that guarantee timeseries well-behavedness and measurement feasibility (Appendix B), selects the windows by comparing data and synthetics, no matter what type of phases are involved. Moreover, using an automated procedure is necessary when dealing with a huge amount of seismograms and measurements. After the selection, one can check the resulting windows and manually remove the ones that are still inadequate.

To summarize the selection procedure, following Maggi et al. (125) (2009), we briefly outline the steps of `FLEXWIN` algorithm in what follows (for more details see Maggi et al. (125), 2009).

- *Stage A: pre-processing.* Preliminary processing operations can be performed on data and synthetics, but in our case we already did this before using `FLEXWIN`. Then, considering some signal-to-noise criteria (r_P and r_A in Table B.1), seismograms dominated by noise are rejected. Finally, the short-term average/long-term average ratio (STA:LTA), a waveform derived from the envelope of the synthetic seismogram, is defined for each synthetic.

- *Stage B: definition of preliminary measurement windows.* The preliminary candidate windows are centered around the local maxima of the STA:LTA waveforms.

[1] freely available at http://www.geodynamics.org/svn/cig/seismo/3D/ADJOINT_TOMO/flexwin/.

These maxima correspond in position and width to the seismic phases in the synthetic seismograms. In order to be considered acceptable, the maxima should lie above a user-defined water level ($w_E(t)$ in Table B.1).

- *Stage C: rejection of preliminary windows based on the content of the synthetic seismogram alone.* Among the set of candidate windows from Stage B, one starts rejecting windows based upon the shape of STA:LTA waveforms. The objective is to retain windows that contain well-developed seismic phases or groups of phases of the synthetics (the parameters used in this stage are c_{0-4} in Table B.1).

- *Stage D: rejection of preliminary windows based on the differences between observed and synthetic seismograms.* Within the windows resulting from stage C, one evaluates four quantities that quantify the well-behavedness of the data and the similarity to the synthetics: (1) signal-to-noise ratio for the data, (2) normalized cross-correlation CC between data and synthetics, (3) cross-correlation time lag $\Delta\tau$, and (4) amplitude ratio $\Delta\ln A$. These four quantities should belong to specific ranges based upon user-defined criteria ($r_0(t)$, $CC_0(t)$, $\Delta\tau_0(t)$, $\Delta\ln A_0(t)$, $\Delta\tau_{ref}$, $\Delta\ln A_{ref}$ in Table B.1). Only the windows that satisfy these conditions are retained.

- *Stage E: resolution of preliminary window overlaps.* Considering the windows from stage D, the final set is chosen so that it contains time windows that do not overlap with each other and that cover as much of the seismogram as possible (the considered criteria are w_{CC}, w_{len} and w_{nwin} in Table B.1).

The windows resulting from the five stages of `FLEXWIN` are those within which the measurements should be estimated (section 4.6.1). They represent the only part of the time series involved in the inversion. Nevertheless, successive iterations of the inversion procedure tend to improve the fit between data and synthetics (e.g., Figs. 4.26 - 4.37). This will cause an extension of the windows and/or an increase of their number, aiming to the ideal condition where the whole waveforms are fitted.

In Appendix B is reported a description of the user parameters set in `FLEXWIN`. Depending on the specific considered scenario, the user parameters should assume different values. Thus, we have performed preliminary tests in order to find the suitable setting for our problem. In particular, we allow both body and surface waves to be

selected in the seismograms (by appropriately tuning the water level in Stage B above). Moreover, conservative criteria to asses timeseries similarity are preferred. The chosen parameter values for central Italy tomographic inversion are reported in Table B.2 for both of the considered period ranges, (2-20) s and (6-20) s.

4.6 Measurements and misfit functions

4.6.1 Multitaper traveltime (MT-TT) misfit function and adjoint source

Once the windows have been selected, the measurements need to be estimated. In order to quantify the misfit between data and synthetics for the adjoint traveltime tomography of central Italy, we choose to measure the frequency-dependent multitaper difference between observed and theoretical traveltimes (Thomson (194), 1982; Park et al. (153), 1987; Percival & Walden (159), 1993; Laske & Masters (112), 1996; Ekström et al. (67), 1997; Zhou et al. (212, 213), 2004, 2005). Thus, for each time window p selected in a single seismogram for a given event s and a given category c (i.e., choosing for example a ground motion component and/or a period range, see section 2.3.4.2), we estimate a *multitaper traveltime (MT-TT) measurement* as a function of frequency ω

$$\Delta\tau_p\ (\omega, \mathbf{m}) = \tau_p^{obs}(\omega) - \tau_p(\omega, \mathbf{m})\,, \tag{4.6}$$

where $\mathbf{m}$ is the model vector. The corresponding *multitaper traveltime misfit function* for a given time window p per event s and category c is given by (e.g., Tape et al. (189), 2010)

$$F_p(\mathbf{m}) \;=\; \frac{1}{H_p}\int_{-\infty}^{\infty} h_p(\omega)\left[\frac{\Delta\tau_p\ (\omega, \mathbf{m})}{\sigma_p(\omega)}\right]^2 \mathrm{d}\omega\,, \tag{4.7}$$

where $\Delta\tau_p(\omega, \mathbf{m})$ is given by 4.6 and $\sigma_p(\omega)$ is the frequency-dependent uncertainty associated with the traveltime measurement. Moreover, $h_p(\omega)$ denotes a windowing function in the frequency domain and H_p is the corresponding normalization factor defined as

$$H_p = \int_{-\infty}^{\infty} h_p(\omega)\,\mathrm{d}\omega\,. \tag{4.8}$$

One can introduce a quantity $w_p(\omega)$ that involves the variance, the windowing filter and the normalization factor:

$$w_p(\omega) = \frac{h_p(\omega)}{H_p\,\sigma_p^2(\omega)}\,, \tag{4.9}$$

thus, the variation of the misfit function 4.7 is given by

$$\delta F_p(\mathbf{m}) = -2 \int_{-\infty}^{\infty} w_p(\omega)\, \Delta\tau_p(\omega, \mathbf{m})\, \delta\tau_p(\omega, \mathbf{m})\, \mathrm{d}\omega \,, \tag{4.10}$$

where $\delta\tau_p(\omega, \mathbf{m})$ represents the frequency-dependent traveltime perturbation due to changes in the model parameters.

Note that, for frequency-independent measurements, $\Delta\tau_p(\omega, \mathbf{m})$ in 4.6 reduces to the cross-correlation traveltime (CC-TT) difference

$$\Delta T_p(\mathbf{m}) = T_p^{obs} - T_p(\mathbf{m}) \,, \tag{4.11}$$

previously introduced (eq. 2.75 with $r = p$ index of the window). Thus, the misfit function 4.7 becomes (by using 4.8)

$$F_p^{CC}(\mathbf{m}) = \left(\frac{\Delta T_p(\mathbf{m})}{\sigma_p} \right)^2 \,, \tag{4.12}$$

that is the cross-correlation traveltime misfit function in equation 2.127 for a single window p.

In the above formulations, as required, the measurements and the misfit functions are defined within a specific time window p of a given seismogram. Thus, the windowing function to be applied to data and synthetics is a crucial choice, since it affects the estimate of the measurements and, then, of the corresponding misfit functions and adjoint sources. As anticipated, we decide to use a multitaper measurement technique (Thomson (194), 1982). This method uses a set of orthogonal tapers t_j, called *prolate spheroidal eigentapers* (Slepian (179), 1978), to window the data and the synthetics. The advantage is that these multi-tapers are all optimally concentrated within a small window in the frequency domain, and they do not have side-lobes. Thus, using this technique, one can minimize both the spectral leakage, caused for example by using windowing functions in the frequency domain with side-lobes (as sinc-functions), and also the bias in the measurements, caused for example by cosine windowing functions. Moreover, since we have multiple estimates for each measurement in the frequency domain, one for each taper, one can also calculate the uncertainty associated to the estimate. Finally, with respect to a cross-correlation measurement, a multitaper technique accounts for the traveltime difference between data and synthetics as a function of frequency in the chosen bandwidth, thus a multitaper measurement, when possible, represents a more appropriate measure of the misfit between two timeseries.

By using the above multitapers t_j to window the data d and the synthetics s, for each t_j one obtains a tapered version of the two timeseries, that, in the time domain, is given by

$$d_j(t) = d(t)\, t_j(t)\,, \qquad s_j(t) = s(t)\, t_j(t)\,. \tag{4.13}$$

In the frequency domain, the windowing operation corresponds to a convolution between the Fourier transform (i.e., the spectra) of the timeseries, $d(\omega)$ and $s(\omega)$, and the Fourier transform of the taper, $t_j(\omega)$, that is

$$d_j(\omega) = d(\omega) \,*\, t_j(\omega)\,, \qquad s_j(\omega) = s(\omega) \,*\, t_j(\omega)\,. \tag{4.14}$$

In order to determine the adjoint sources for multitaper traveltime measurements, one can follow the formulation in Tape (186) (2009, Appendix C) (see details therein). The variation of the misfit function in 4.10 needs to be expressed in terms of the perturbed displacement δs, as previously done for the cross-correlation traveltime misfit function (eqs. 2.76-2.77). Based upon Tape (186) (2009), the traveltime perturbation $\delta\tau(\omega)$ in 4.10 can be written

$$\delta\tau(\omega) = -\frac{1}{\omega}\, \mathrm{Im}\left(\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*} \right)\,, \tag{4.15}$$

where s_j is the tapered synthetic in 4.14 and δs_j the corresponding tapered, perturbed synthetic. Note that in 4.15 and in the following equations we omit, for sake of simplicity, the index p of the window, the index i of the ground motion component (present in eq. 2.77) and the dependence on $\mathbf{m}$ and $\mathbf{x}$. By substituting 4.15 in 4.10, after some algebra (see, for details, Tape (186), 2009), one obtains

$$\delta F = 2\,\mathrm{Re}\left[\sum_j \int_{-\infty}^{\infty} w(\omega)\, \Delta\tau(\omega)\; p_j(\omega)\, \delta s_j^*(\omega)\, \mathrm{d}\omega \right]\,, \tag{4.16}$$

where the function $p_j(\omega)$ in the frequency domain and for the j-th taper is

$$p_j(\omega) = \frac{i\omega s_j}{\sum_{k'} (i\omega s_{k'})\, (i\omega s_{k'})^*}\,. \tag{4.17}$$

By transforming expression 4.16 into the time domain [1], one can write the variation of the mutitaper traveltime misfit function in terms of the perturbed time series $\delta s(t)$ as

$$\delta F = \int_{-\infty}^{\infty} f^\dagger(t)\, \delta s(t)\, \mathrm{d}t\,, \tag{4.18}$$

[1] by using the Plancherel's theorem $\int_{-\infty}^{\infty} f(\omega)\, g^*(\omega)\, \mathrm{d}\omega = 2\pi \int_{-\infty}^{\infty} f(t)\, g^*(t)\, \mathrm{d}t$ (see again Tape (186), 2009)

and, thus, the *multitaper traveltime adjoint source* $f^{\dagger}(t)$, for a given window selected in a single seismogram, is expressed as

$$f^{\dagger}(t) = \sum_{j} t_j(t) \left[4\pi \, w(t) * \Delta\tau(t) * p_j(t) \right] . \tag{4.19}$$

It is worth noting that $p_j(t)$ in 4.19 is the time-domain version of $p_j(\omega)$ in 4.17, i.e. its inverse Fourier transform, and it contains the first time derivative of the tapered synthetic, $\partial_t s_j(t)$. Thus, by comparing equations 4.19 and 2.96 (written for a single value of r and weighted by σ), one can recognize the similarities between multitaper and cross-correlation traveltime adjoint sources. Next, by following Tromp et al. (197) (2005) as in section 2.3.3, the adjoint field can be defined in terms of the adjoint source and the variation of the multitaper traveltime misfit function can be written again as a volume integral involving 3D sensitivity kernels (as, e.g., eq. 2.89).

For the adjoint tomography of central Italy, after selecting the windows suitable for measurements through `FLEXWIN` (section 4.5.2), we use the code `measure_adj`[1] to estimate, within these windows, the measurements and, then, the corresponding misfit functions and adjoint sources at each iteration. The code allows one to choose among different types of measurements and misfit functions to be evaluated, and in our case, as mentioned, we ask for calculation of multitaper traveltime measurements. `measure_adj` performs an initial check of the windows selected by `FLEXWIN` and, accounting for some user-defined parameters (Appendix C), it executes two distinct calculations described in what follows:

- when the windows satisfy suitable conditions based upon the user parameters, the code applies the multitaper technique to window data and synthetics, and estimates the MT-TT differences 4.6 as a function of frequency. Then, it calculates, for each window p, the MT-TT misfit function in 4.7 and the MT-TT adjoint source in 4.19. (See the last three equations in Table 4.1).

- whenever the time window does not satisfy the user parameter conditions (e.g., the window is too narrow), the MT-TT measurement on this window is reverted to the CC-TT measurement 4.11. In this case the CC-TT misfit function 4.12 and the CC-TT adjoint source 2.96 (weighted by σ) are evaluated for each p. (See the first three equations in Table 4.1).

[1] freely available at http://www.geodynamics.org/svn/cig/seismo/3D/ADJOINT_TOMO/measure_adj/

Table 4.1 summarizes the measurements, misfit functions and adjoint sources (for a window p) considered for central Italy adjoint tomography.

Note that the uncertainties σ_p associated to each measurement appear in the denominator of both the misfit functions and the adjoint sources (Tab. 4.1). Thus, in order to avoid singularities for $\sigma_p = 0$ (a perfect measurement), one needs to define a minimum threshold value σ_0 for σ_p. Following Tape et al. (189) (2010), in our tomographic inversion we choose

$$\sigma_0 = 1.0 \text{ s}\,, \tag{4.20}$$

based upon the range of M_w (Fig. 4.3 (c)) and the uncertainties associated to the source parameters. The code `measure_adj` calculates the value of the uncertainties on the measurements using the formulation in Tape et al. (189) (2010, Appendix A) and it assigns to σ_p the value of σ_0, whenever $\sigma_p < \sigma_0$.

4.6.2 Overall misfit

In section 2.3.4.2 we introduced the misfit function F_c for a given category c and the overall misfit function F. In the adjoint inversion for central Italy, as previously introduced (section 4.5.1), we work with three components (radial, vertical and transverse) of data and synthetics within two period ranges, 2-20 s and 6-20 s. Each pair [component - period band] represents a category, thus we consider a total of 6 categories. At each iteration of the inversion procedure we are interested in monitoring the behaviour of the six misfit functions F_c per category and of the total misfit F; they all are expected to decrease for successive iterations, while improving the model.

Considering the multitaper traveltime misfit function in 4.7, now labeled as F_p^{MT}, the MT-TT misfit function per event s and category c (see eq. 2.126) is obtained by summing over all the time windows N_{sc}^{MT} suitable for a MT measurement (given s and c) and normalizing by this number, that is (see Table 4.1)

$$F_{sc}^{MT} = \frac{1}{N_{sc}^{MT}} \sum_{p=1}^{N_{sc}^{MT}} F_p^{MT} = \frac{1}{N_{sc}^{MT}} \sum_{p=1}^{N_{sc}^{MT}} \frac{1}{H_p} \int_{-\infty}^{\infty} h_p(\omega) \left[\frac{\Delta\tau_p\ (\omega, \mathbf{m})}{\sigma_p(\omega)} \right]^2 \mathrm{d}\omega\,. \tag{4.21}$$

When cross-correlation traveltime measurements are estimated, one sums and normalizes by considering the windows for which a CC-TT measurement has been made. Thus,

	Cross-correlation Traveltime (CC-TT)
Measurement	$\Delta T(\mathbf{m}) = T^{obs} - T(\mathbf{m})$
Misfit Function	$F^{CC}(\mathbf{m}) = \left(\frac{\Delta T(\mathbf{m})}{\sigma}\right)^2$
Adjoint Source	$f^{\dagger}(\mathbf{x},t) = -\,\Delta T\ \frac{2\,w(T-t)}{\sigma^2\,M}\,\partial_t s(\mathbf{x}_r, T-t)\,\delta(\mathbf{x}-\mathbf{x}_r)$
	Multitaper Traveltime (MT-TT)
Measurement	$\Delta\tau(\omega,\mathbf{m}) = \tau^{obs}(\omega) - \tau(\omega,\mathbf{m})$
Misfit Function	$F^{MT}(\mathbf{m}) = \frac{1}{H}\int_{-\infty}^{\infty} h(\omega)\left[\frac{\Delta\tau(\omega,\mathbf{m})}{\sigma(\omega)}\right]^2 \mathrm{d}\omega$
Adjoint Source	$f^{\dagger}(\mathbf{x},t) = \sum_j\ t_j(t)\,[4\pi\,w(t) *\ \Delta\tau(t)\ *\ p_j(\mathbf{x}_r,t)]\,\delta(\mathbf{x}-\mathbf{x}_r)$

Table 4.1: Measurements, misfit functions and adjoint sources for a time window p - Measurements, misfit functions and adjoint sources used in the adjoint tomographic inversion of central Italy. They refer to a single window p, a single component i, a single receiver $\mathbf{x}_r$ and a single event, but indexes p and i have been neglected. Note that both CC and MT adjoint sources contain the measurement uncertainty σ in the denominator, since $w(t)$ in the MT adjoint source is the time-domain version of 4.9. The model vector is denoted by $\mathbf{m}$, the superscript *obs* denotes the observed traveltime, and $p_j(t)$ is the time-domain version of $p_j(\omega)$ in 4.17.

the CC-TT misfit function per event s and category c, labeled as F_{sc}^{CC}, is given by (see eq. 2.127 and Table 4.1)

$$F_{sc}^{CC} = \frac{1}{N_{sc}^{CC}} \sum_{p=1}^{N_{sc}^{CC}} F_p^{CC} = \frac{1}{N_{sc}^{CC}} \sum_{p=1}^{N_{sc}^{CC}} \left(\frac{\Delta T_p}{\sigma_p} \right)^2 . \tag{4.22}$$

In our case, both the MT-TT and the CC-TT differences are measured, depending on whether the windows satisfy some given conditions in `measure_adj` (Appendix C). Thus, we define the total number of windows N_{sc} per event s and category c as

$$N_{sc} = N_{sc}^{MT} + N_{sc}^{CC} , \tag{4.23}$$

and the general misfit function per event s and category c, considered in this work, is given by

$$F_{sc} = \frac{1}{N_{sc}} \left[\sum_{p=1}^{N_{sc}^{MT}} F_p^{MT} + \sum_{p=1}^{N_{sc}^{CC}} F_p^{CC} \right] . \tag{4.24}$$

Finally, the misfit functions for each of the six categories are (from eq. 2.128)

$$F_c = \frac{1}{N_s} \sum_{s=1}^{N_s} F_{sc} , \tag{4.25}$$

where each event has a similar weight ($1/N_s$), with N_s the total number of considered events. Then, the overall misfit function is (from eq. 2.129)

$$F = \frac{1}{6} \sum_{c=1}^{6} F_c , \tag{4.26}$$

where we force an equal weight for each category and we consider a total of 6 categories.

4.7 Inversion procedure

Summarizing the key points of section 2.3.4.1, the iterative inversion steepest descent procedure starts from the initial velocity model $\mathbf{m}^0$ (section 4.2.1). Based upon this model, together with the source representations (section 4.3), one can construct the corresponding synthetic waveforms by using SPECFEM3D. Then, the misfit function between data and synthetics for the initial model, $F(\mathbf{m}^0)$, can be estimated (Tab. 4.1

and eqs. 4.21 - 4.26). At this point, one needs to calculate the initial gradient $\mathbf{g}^0$ of the misfit function

$$\mathbf{g}^0 = \mathbf{g}(\mathbf{m}^0) = \left.\frac{\partial F}{\partial \mathbf{m}}\right|_{\mathbf{m}^0} , \tag{4.27}$$

and this will be used to obtain the first updated model $\mathbf{m}^1$ by (eq. 2.124)

$$\mathbf{m}^1 \,=\, \mathbf{m}^0 \,+\, \alpha\, \mathbf{p}^0 \,, \tag{4.28}$$

where α is the step length and $\mathbf{p}^0$ the initial steepest descent vector (eq. 2.123)

$$\mathbf{p}^0 \,=\, -\,\mathbf{g}^0 \,. \tag{4.29}$$

Analogous equations will hold for the next iterations by substituting 1 with n and 0 with $n-1$ (eq. 2.125 and section 4.9).

In what follows, I will present the construction of the initial gradient of the misfit function for central Italy tomography, based upon the adjoint method, i.e., obtaining the gradient by calculating the adjoint misfit kernels K^0 for the initial model (eq. 2.102)

$$g_k^0 = \left.\frac{\partial F}{\partial m_k}\right|_{\mathbf{m}^0} = \int_V K^0\, B_k \, \mathrm{d}^3\mathbf{x} \qquad k = 1, ..., M \,. \tag{4.30}$$

The regularizations applied to this gradient are discussed in section 4.7.2, and in section 4.7.3 I will describe in details the line search performed to find the step length α for the first model update in equation 4.28. The results at the first iteration are shown in paragraph 4.8. Then, in section 4.9 the following improved models, until the last one we obtained, $\mathbf{m}^5$, are presented, considering that the calculation of the misfit kernels, the regularizations and the steplength test need to be performed at each iteration.

4.7.1 Event kernels and gradient of the misfit function

In order to construct the misfit kernels, one needs to calculate the traveltime event kernels for all the considered events from the interaction between the corresponding regular and adjoint wavefields.

By using the code `measure_adj` (section 4.6.1), the adjoint sources for each window selected in the seismograms for every event have been estimated in both of the considered period ranges, (2-20) s and (6-20) s (Tab. 4.1). A combined adjoint source for each receiver component (per given event) is obtained by summing those in the two

period bands, and this is passed as input to SPECFEM3D. Thus, using the spectral-element code, for each event, one performs a forward and an adjoint simulation, as described in section 2.4. The adjoint sources at all the receivers of an event act simultaneously to generate the adjoint field for the given event (eq. 2.95). The interaction with the corresponding back-reconstructed forward field results in the event kernel for the specific earthquake. Considering the parametrization of the initial velocity model for central Italy (section 4.2.1), the kernels calculated for each event are the compressional wavespeed kernel K^s_α and the shear wavespeed kernel K^s_β given by (eqs. 2.91 - 2.92 and 2.87)

$$K^s_\alpha = 2\left(\frac{\kappa + \frac{4}{3}\mu}{\kappa}\right) K^s_\kappa\,, \qquad K^s_\beta = 2\left(K^s_\mu - \frac{4}{3}\frac{\mu}{\kappa}\,K^s_\kappa\right)\,, \tag{4.31}$$

where

$$K^s_\mu = -2\mu \int_0^T \mathbf{D}^\dagger(T-t) : \mathbf{D}(t)\,\mathrm{d}t\,, \qquad K^s_\kappa = -\kappa \int_0^T \left[\nabla\cdot\mathbf{s}^\dagger(T-t)\right]\left[\nabla\cdot\mathbf{s}(t)\right]\,\mathrm{d}t\,. \tag{4.32}$$

The superscript s highlights that we refer to a single event, but it has been neglected for $\mathbf{s}$, $\mathbf{s}^\dagger$, $\mathbf{D}$ and $\mathbf{D}^\dagger$ to avoid cluttering. Note that, in our tomographic inversion, the adjoint field $\mathbf{s}^\dagger$ in 4.31 - 4.32 ($\mathbf{D}^\dagger$ is given by 2.62) is generated by the (MT or CC) traveltime adjoint sources in Table 4.1 summed over all the windows for all the receivers per given event.

Considering the event kernels in 4.31, the traveltime misfit kernels for the initial model are obtained by summing these event kernels over all the considered earthquakes, i.e. (from eq. 2.121)

$$K_\alpha = \sum_{s=1}^{N_s} K^s_\alpha\,, \qquad K_\beta = \sum_{s=1}^{N_s} K^s_\beta\,, \tag{4.33}$$

where for the initial model $\mathbf{m}^0$ we sum over all the 63 events considered in the inversion (section 4.3), i.e., $N_s = 63$.

The event kernels for different events can be different from each other. They also differ from the corresponding misfit kernel, that is their sum and thus it represents the total sensitivity of the waveforms to model perturbations. Figure 4.4 shows the kernel for two of the 63 considered events and the total misfit kernel for the initial model $\mathbf{m}^0$, highlighting their different sensitivity.

The misfit kernels in 4.33 give us directly the initial gradient $\mathbf{g}^0$ of the misfit function $F(\mathbf{m}^0)$, because in an adjoint approach equation 4.30 must hold. Then, since we use a spectral-element method, the basis functions B_k in 4.30 can be defined based on the Lagrange polynomial basis embedded in the numerical code. Thus, as shown in section 2.3.4, the gradient of the misfit function is practically constructed from the misfit kernels defined at all the GLL points that discretize the SEM volume, i.e. (eq. 2.120)

$$g_k^0 = K_k^0 \sqrt{V_k} \qquad k = 1, ..., N_{GLL}\,, \tag{4.34}$$

where V_k is the volume associated with the k-th GLL point (eq. 2.117) [1]. The discrete misfit kernels in 4.34 are exactly the ones produced by SPECFEM3D and they are shown in Figure 4.5 for the initial model $\mathbf{m}^0$.

One should be aware in interpreting the kernels, since they are gradients of a misfit function at a particular iteration. However, they in general evidence where the current model differs from the real structure and they give an idea of the direction along which the model should be modified to reduce the misfit. Given the convention on the traveltime measurements (section 2.3.3), positive-blue values of the kernels correspond to positive measurements, suggesting that the velocity of our model should be decreased to fit the data. On the contrary, negative-red values indicate that velocities need to be increased. Looking at Figure 4.5, our initial velocity model for central Italy seems to be too fast almost everywhere in the considered volume. Thus, we expect that the first model update will in general decrease the velocity values.

4.7.2 Preconditioning and Smoothing

The misfit kernels obtained by simply summing the event kernels (Fig. 4.5) may contain some spurious amplitudes near sources and receivers. Moreover, they can show structures with a scalelength not properly resolved by the simulations. In order to reduce the fictitious effects, and not to include finer features in the updated model until they

[1] The superscript 0 emphasizes that we refer to the initial model $\mathbf{m}^0$, but in equations 4.31 - 4.33 it has been neglected to avoid clutters. Moreover, analogous relationships as 4.31 - 4.34 will hold for every iteration (section 4.9). In general, given our parametrization, the components of the gradient vector are $\mathbf{g} = [K_{\alpha\,1}\sqrt{V_1}, ..., K_{\alpha\,k}\sqrt{V_k}, ..., K_{\alpha\,N_{GLL}}\sqrt{V_{N_{GLL}}}, K_{\beta\,1}\sqrt{V_1}, ..., K_{\beta\,k}\sqrt{V_k}, ..., K_{\beta\,N_{GLL}}\sqrt{V_{N_{GLL}}}]^T$ (e.g., Tape et al. (189), 2010).

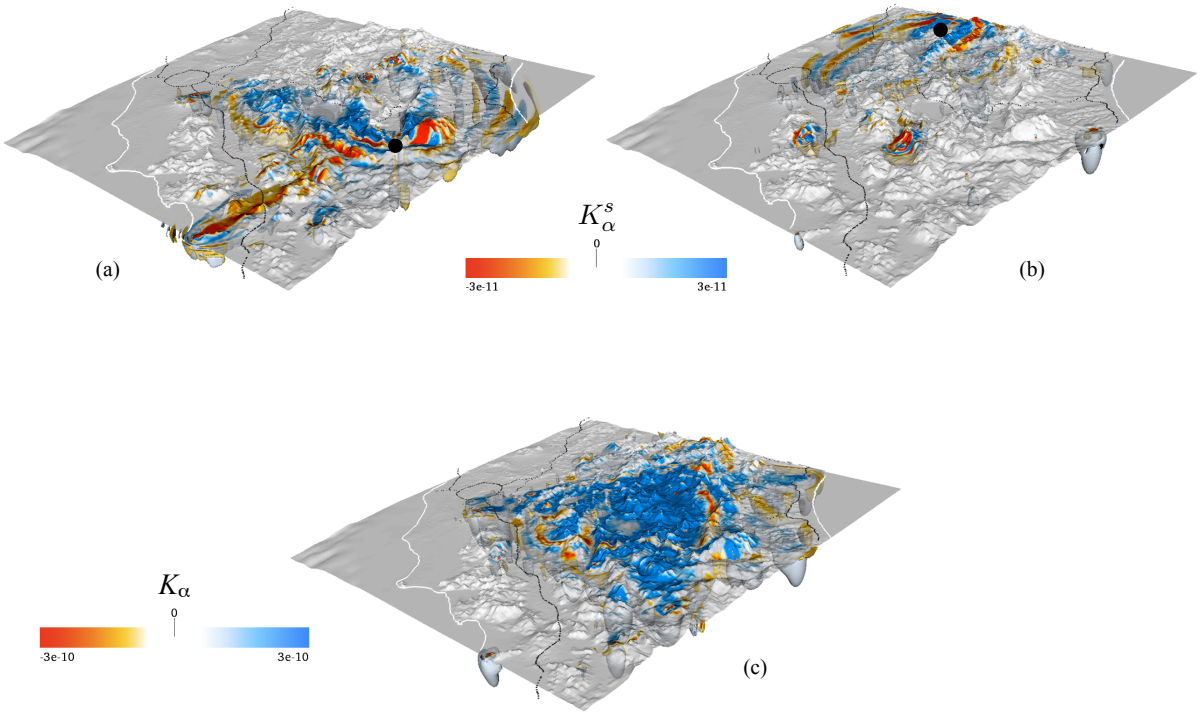


Figure 4.4: Event kernels and misfit kernel - (a) and (b) represent the event kernels K^s_α (eq. 4.31) for the initial model $\mathbf{m}^0$ and for two of the considered events, indicated by the black dots: (a) event of March 17th 2009 at 01:12 UTC, (b) event of August 28th 2010 at 07:08 UTC. (c) Misfit kernel K_α (eq. 4.33) for model $\mathbf{m}^0$, obtained by summing the event kernels of all the 63 considered events. K^s_α and K_α are in m^{-3}.

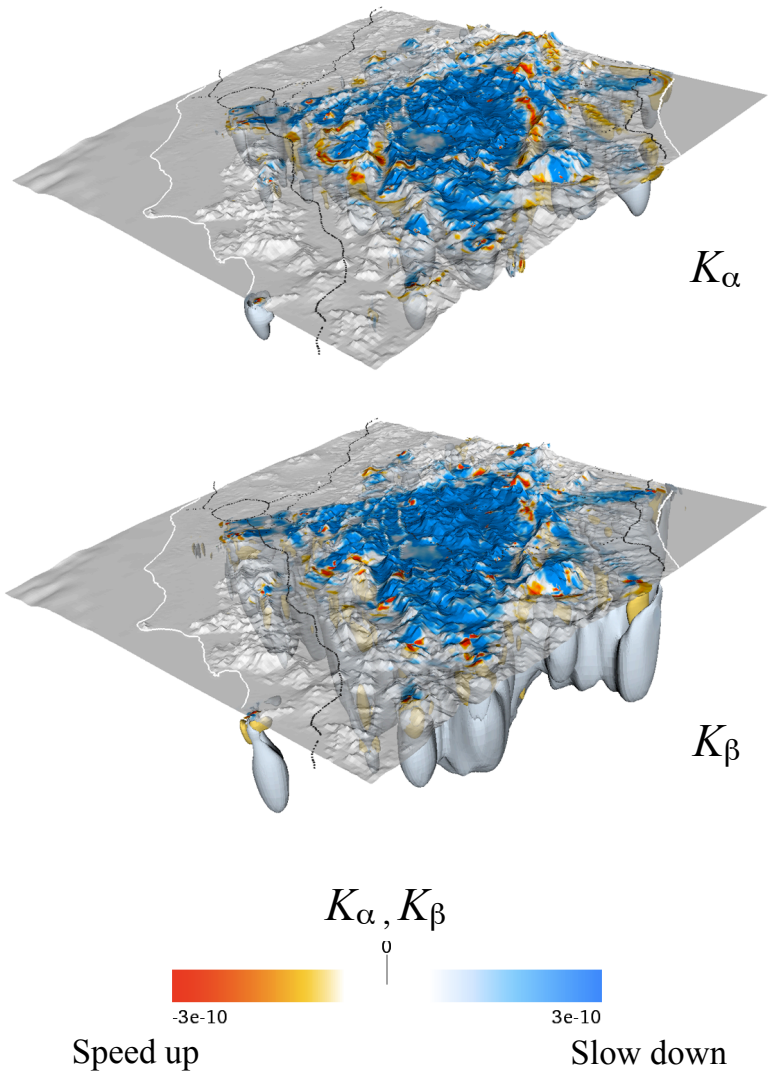


Figure 4.5: Misfit kernels for the initial model $\mathbf{m}^0$ - Misfit kernels K_α and K_β (eq. 4.33) for the initial velocity model $\mathbf{m}^0$ for central Italy (section 4.2.1). The positive-blue values indicate that the initial model is too fast with respect to the real structure, thus one expects that the model update will decrease the velocities in $\mathbf{m}^0$. Viceversa for negative-red values. The kernels are in m^{-3} and are the sums of the event kernels for all the 63 events. The simulated volume in central Italy is viewed from SE; coastlines are white, highways are black.

will be effectively resolved, we need to apply some regularizations to the original misfit kernels. The regularizations can also help to increase the convergence of the procedure.

Following section 2.3.5, we firstly perform a preconditioning of the misfit kernels 4.33 for the initial model $\mathbf{m}^0$, by using the total preconditioner P summed over all the events N_s (eqs. 2.130 - 2.131) [1]

$$P(\mathbf{x}) \,=\, \sum_{s=1}^{N_s} P^s(\mathbf{x}) \,=\, \sum_{s=1}^{N_s} \int_0^T \partial_t^2 \mathbf{s}(\mathbf{x},t) \cdot \partial_t^2 \mathbf{s}^\dagger(\mathbf{x}, T-t)\, \mathrm{d}t\,. \tag{4.35}$$

Thus, by using the inverse of matrix 4.35, the preconditioned versions of the initial misfit kernels in 4.33 are given by (from eq. 2.132)

$$K_\alpha^{PR} \,=\, P^{-1}\, K_\alpha\,, \qquad K_\beta^{PR} \,=\, P^{-1}\, K_\beta\,. \tag{4.36}$$

Note that, as introduced in section 2.3.5, the preconditioners P^s in 4.35 for each event are calculated by SPECFEM3D during the calculation of the event kernels themselves.

To regularize the inversion one can also apply a smoothing to the initial misfit kernels by convolving them with a 3D Gaussian $G(\mathbf{x})$ (section 2.3.5). Thus, the smoothed versions of the misfit kernels in 4.33 for model $\mathbf{m}^0$ are given by (from eqs. 2.134 - 2.135)

$$K_\alpha^{SM} \,=\, K_\alpha * G\,, \qquad K_\beta^{SM} \,=\, K_\beta * G\,. \tag{4.37}$$

The smoothing aims to remove from the kernels features with scalelengths smaller than the shortest wavelengths resolved in the problem. Thus, the scalelengths of the smoothing in the horizontal and vertical directions, i.e., the half-widths σ of the 3D Gaussian (eq. 2.136), should be chosen accordingly. In our tomographic inversion for central Italy, considering the size of the mesh elements ($\sim$1 km on the top, section 4.2.2), the estimated accuracy of the simulations ($\sim$ 0.5-1 Hz) and the scalelength of kernel variations, we use a quite conservative Gaussian smoothing by setting

$$\sigma_x \,=\, \sigma_y \,= 5\,\mathrm{km} \qquad \sigma_z \,= 1\,\mathrm{km}\,. \tag{4.38}$$

This choice is based upon some empirical tests that we performed for different values of smoothing. There is always a trade-off between not removing too much signal and minimizing the spurious features. Note that σ_z is lower than σ_x and σ_y, because the scalelength of the problem and the kernel sensibility are lower in the vertical direction

[1] Superscript 0 referring to $\mathbf{m}^0$ has been neglected, as superscript s for $\mathbf{s}$ and $\mathbf{s}^\dagger$.

with respect to the horizontal ones. We prefer to use the same smoothing 4.38 for all the iterations considered in our work. A less conservative choice would be to reduce the smoothing for successive iterations, since the problem resolution increases. This could accelerate the minimization procedure, but may also cause a step back in following iterations, if spurious effects are introduced, justifying our choice to maintain the same smoothing at least for first iterations.

Finally, in general, one could apply both the preconditioning and the smoothing, for example, by convolving the 3D Gaussian with the preconditioned misfit kernels in 4.36. One, thus, obtains the preconditioned, smoothed misfit kernels

$$K_{\alpha}^{PR-SM} = K_{\alpha}^{PR} * G\,, \qquad K_{\beta}^{PR-SM} = K_{\beta}^{PR} * G\,. \tag{4.39}$$

To choose which one of the misfit kernels 4.37 or 4.39 should represent the gradient of the misfit function via equation 4.34, we performed some preliminary tests of their effects on the misfit function values (see section 4.7.3).

For the initial model $\mathbf{m}^0$ all the kernels K_α in 4.33, 4.36, 4.37 and 4.39 are reported, as an example, in Figure 4.6 to show the behaviour of the different regularizations. One can observe how the preconditioning constrains the sensibility of the kernels and how the smoothing tends to remove the shorter scalelength structures likely unresolved. However, it is worth noting that the dominant feature of suggesting a decreasing of the initial model velocities is preserved by all the kernels.

4.7.3 Steplength test

After calculating the initial gradient of the misfit function, in order to use equation 4.28 to find the first improved model $\mathbf{m}^1$, one needs to determine the step length α. This value quantifies the percentage of the perturbation that should be applied to the initial model in order to minimize the misfit function.

As explained in section 2.3.4.1, the step length can be estimated for each iteration by performing a line search for different values of α. For each attempted α, one needs to run a full-waveform simulation for every event. Thus, to reduce the computational costs, we started by selecting a subset of representative earthquakes among the 63 initially considered (section 4.3). We choose 11 events out of the 63, with a very high number of measurements in both of the considered period bands. Moreover, these earthquakes have been selected so that their distribution covers as much of the considered volume

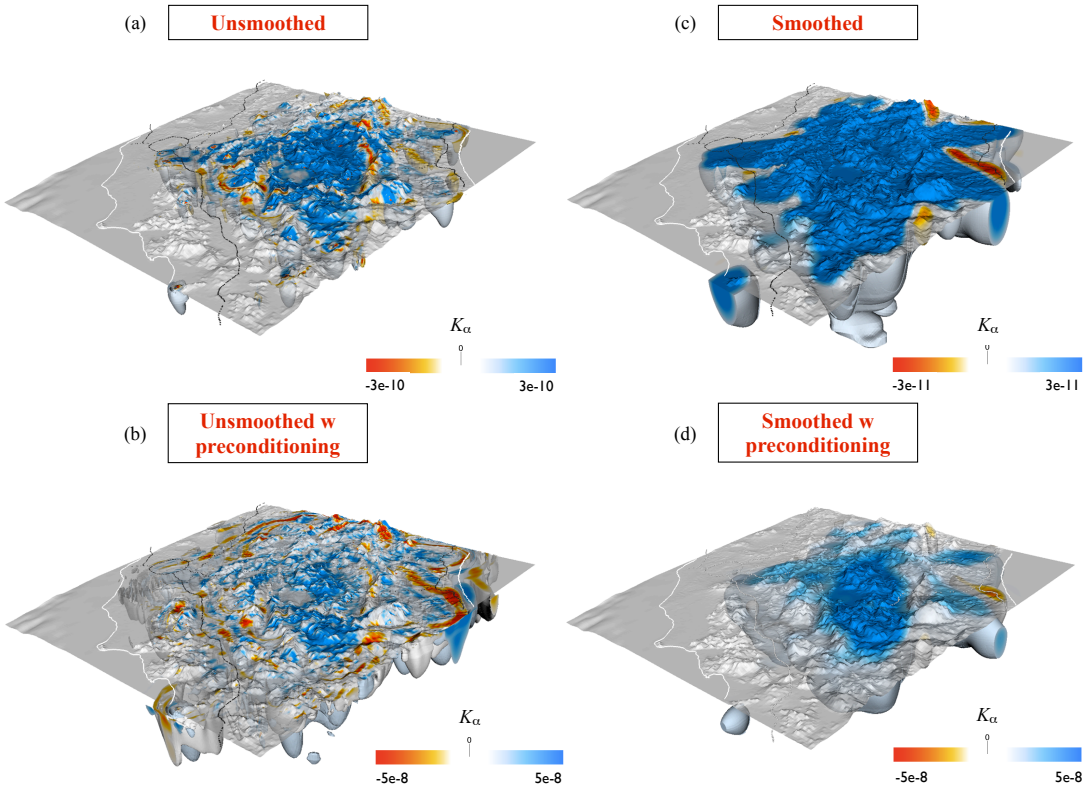


Figure 4.6: Misfit kernel regularizations - (a) Original misfit kernel K_α (eq. 4.33) for the initial model $\mathbf{m}^0$ without applying any regularization. (b) Preconditioned K_α (eq. 4.36) obtained by applying to the kernel in fig.(a) the preconditioner in 4.35. (c) Smoothed K_α (eq. 4.37) obtained by applying to the kernel in fig.(a) a Gaussian smoothing with σ in 4.38. (d) Preconditioned and smoothed K_α (eq. 4.39) obtained by applying to the kernel in fig.(a) the preconditioner in 4.35 and the Gaussian smoothing with σ in 4.38. All the kernels are in m^{-3}, but the scales are different.

as possible. The chosen 11 events are represented in Figure 4.7 (a), together with some statistics on their parameters (Figs. 4.7 (b)-(c)).

Next, by considering the gradient $\mathbf{g}^0$ of the initial misfit function and the corresponding vector $\mathbf{p}^0$ in 4.29, we calculated four different updated model vectors $\mathbf{m}_i^1$ based upon

$$\mathbf{m}_i^1 = \mathbf{m}^0 + \alpha_i \mathbf{p}^0 \qquad i = 1, 2, 3, 4\,, \tag{4.40}$$

where for the initial model the tested values of steplength are $\alpha = (0.01, 0.02, 0.03, 0.04)$[1]. Moreover, only for the initial model, we tested the four model updates both for the smoothed kernels and for the preconditioned-smoothed kernels. Thus, summarizing:

- we calculated the four updated models in 4.40 by constructing the gradient in 4.34 (then $\mathbf{p}^0$ in 4.40) from the smoothed misfit kernels K_α^{SM} and K_β^{SM} in 4.37 for model $\mathbf{m}^0$;
- we also calculated the four updated models in 4.40 by obtaining $\mathbf{g}^0$ and $\mathbf{p}^0$ from the preconditioned-smoothed misfit kernels K_α^{PR-SM} and K_β^{PR-SM} in 4.39 for model $\mathbf{m}^0$.

For each of the tested model vectors $\mathbf{m}_i^1$ (with and without preconditioning) we simulated the corresponding synthetic seismograms with SPECFEM3D.

At this point, comparing the synthetics with the data, one needs to estimate the misfit function for each possible new model; this will allow us to choose the most suitable value of α. To this end, we prefer to compare the behaviour of the misfit within the same time windows. Thus, we do not rerun `FLEXWIN`, but using the windows selected in the seismograms for the initial model $\mathbf{m}^0$, we run only `measure_adj` in order to estimates the new measurements for the tested updated models. Finally, we obtain two curves (one for the case with preconditioning and the other without preconditioning; Fig. 4.7 (d)), that describe the trend of the misfit function varying the value of α.

Analysing the results of the above described steplength test shown in Figure 4.7 (d), one can determine the best update from the initial model $\mathbf{m}^0$ to the new model $\mathbf{m}^1$. The red dots in the picture indicate the misfit behaviour when the smoothed kernels 4.37 are used to estimate $\mathbf{p}^0$ in 4.40; the green dots are for the case with preconditioned

[1]More specifically, we assign to a step factor the values (0.01, 0.02, 0.03, 0.04). Then, to obtain the values of α we divide each step factor value by the maximum of the moduli of the minimum and maximum values of K_α and K_β.

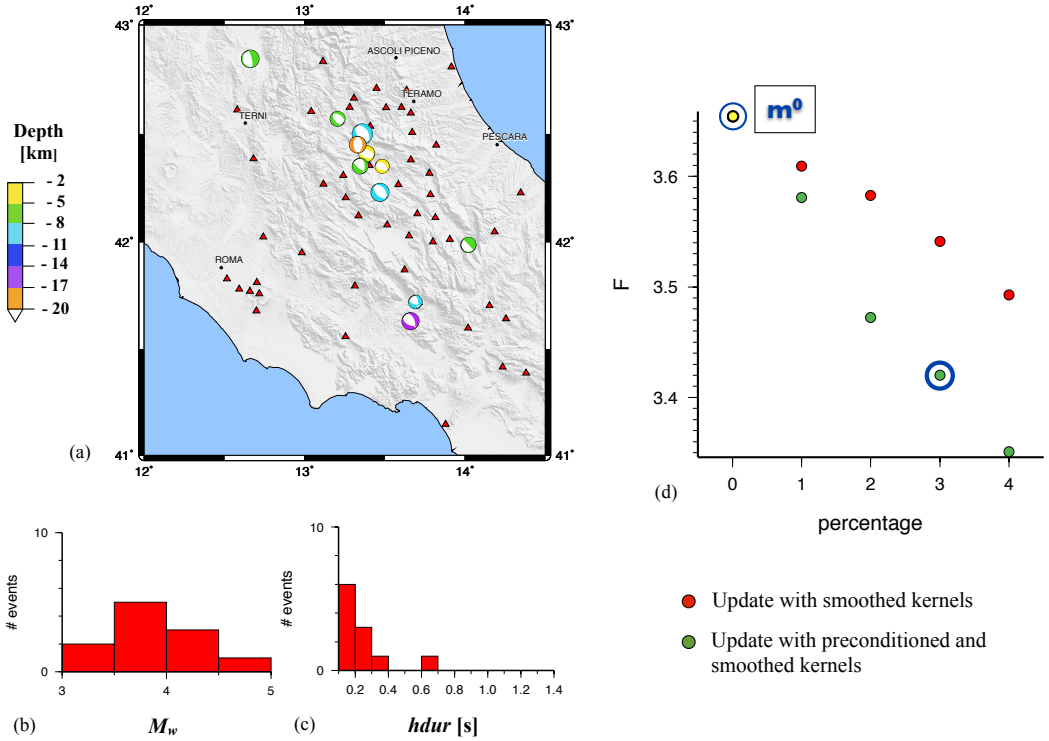


Figure 4.7: Steplength test - (a) The 11 events chosen to perform the steplength test and the considered 52 velocimetric stations (red triangles). The beachballs have dimensions proportional to event magnitudes and colors referred to the depth color scale on the left. (b) Distribution of TDMT moment magnitude M_w for the considered 11 events. (c) Distribution of the half durations *hdur*. (d) Results of the steplength test to determine the update of model $\mathbf{m}^0$. The yellow dot indicates the value of the total traveltime misfit function F (eq. 4.26) for model $\mathbf{m}^0$ (and the 11 events). The red dots indicate the values of F for the 4 possible models $\mathbf{m}_i^1$ obtained by constructing $\mathbf{p}^0$ in 4.40 from the smoothed misfit kernels (eq. 4.37) of the initial model $\mathbf{m}^0$. The green dots indicate the values of F for the 4 possible models $\mathbf{m}_i^1$ obtained by constructing $\mathbf{p}^0$ in 4.40 from the preconditioned-smoothed misfit kernels (eq. 4.39) of model $\mathbf{m}^0$.

and smoothed kernels. In both cases the misfit function deceases by increasing the value of α, as expected. However, one can note how the introduction of the preconditioner, providing additional constraints to the gradient, increases the convergence rate of the minimization algorithm, with respect to the case without preconditioning. Thus, hereafter we prefer to focus our attention only on model updates based upon smoothed, preconditioned misfit kernels, that may help to speed up the inversion procedure.

In order to choose the value of α that finally gives the new model $\mathbf{m}^1$, one should in principle find a minimum of the misfit function test curve. Considering Figure 4.7 (d), since the misfit continues decreasing, the minimum may occur for larger values of the step length not considered in our test. However, working with gradients one should stay in a linear regime. Moreover, applying remarkable perturbations to the starting model could accelerate the inversion procedure, but again one may risk to go too far – the inversion may fall into a local minimum. Thus, we prefer to be conservative and not to consider large updates at least for few iterations. This is also the reason why we finally consider that a suitable value of steplength to obtain the first model update is $\alpha = 0.03$ (corresponding to the circled green dot in Figure 4.7 (d)).

As a consequence, the first improved model $\mathbf{m}^1$ for central Italy is obtained from 4.28, by updating the initial model $\mathbf{m}^0$ with 3% of the negative initial preconditioned-smoothed misfit kernels 4.39. This model will be described in section 4.8. Figure 4.8 shows the preconditioned and smoothed kernels K_α^{PR-SM} and K_β^{PR-SM} (eq. 4.39) for the initial model $\mathbf{m}^0$, used to obtain the new model $\mathbf{m}^1$. As previously discussed (section 4.7.1), the positive-blue values of the kernels in the picture will result in values of v_p and v_s of $\mathbf{m}^1$ decreased with respect to $\mathbf{m}^0$ (section 4.8). Note that the sensitivity in depth of K_β is considerably larger than that of K_α.

It is worth remarking that the steplength test should be performed for each iteration n, to establish which percentage of the kernels at iteration $n-1$ should be used to obtain the corresponding model update. Thus, to test the values of α in order to determine the model $\mathbf{m}^n$ at iteration n, equation 4.40 will be replaced by (see section 4.9)

$$\mathbf{m}_i^n \;=\; \mathbf{m}^{n-1} \;+\; \alpha_i \, \mathbf{p}^{n-1} \qquad i = 1, 2, 3, 4 \,. \tag{4.41}$$

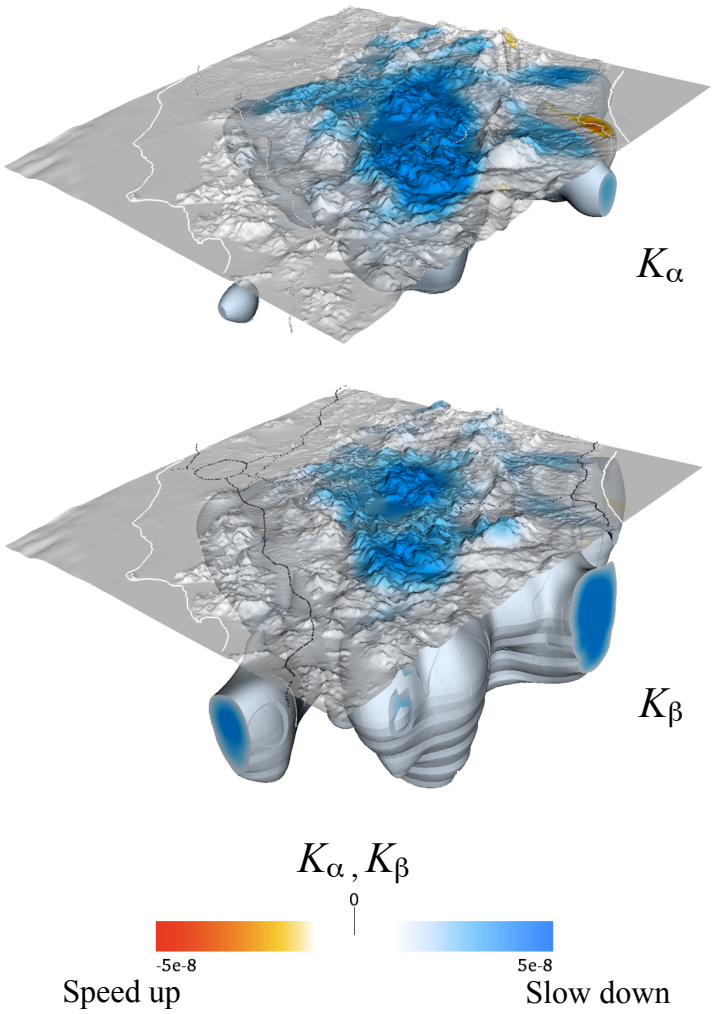


Figure 4.8: Regularized misfit kernels for the initial model $\mathbf{m}^0$ - Preconditioned and smoothed misfit kernels K_α and K_β (eq. 4.39) for the initial model $\mathbf{m}^0$, obtained by applying to the kernels in Figure 4.5 the preconditioner in 4.35 and the Gaussian smoothing with σ in 4.38. The kernels are in m^{-3} and are summed over all the 63 event kernels.

4.8 First iteration updated model $\mathbf{m}^1$

As a result of the steplength test (section 4.7.3), we choose to construct the first updated model $\mathbf{m}^1$ by applying to the initial velocity model $\mathbf{m}^0$ a perturbation equal to 3% of the steepest descent vector. This vector, in turn, has been obtained from the preconditioned and smoothed kernels of $\mathbf{m}^0$ in Figure 4.8.

In order to present the characteristics of the first iteration model, one can starts observing how the velocities have been changed from the initial model $\mathbf{m}^0$. Figure 4.9 shows the normalized variations of v_p and v_s between the new model $\mathbf{m}^1$ and the initial model $\mathbf{m}^0$, i.e.

$$\frac{v(\mathbf{m}^1) - v(\mathbf{m}^0)}{v(\mathbf{m}^0)} \,. \tag{4.42}$$

Most of the variations occur, as expected, around L'Aquila region, since the sources are mainly clustered in this area and thus the kernels have a higher sensibility. The principal result is that the values of the variations, both on the surface and inside the simulated volume, are mostly negative (red in Fig. 4.9). This means that, in general, to obtain the updated model $\mathbf{m}^1$, both the values of v_p and v_s have been decreased with respect to $\mathbf{m}^0$, as suggested by the kernels of the initial model used for the update (Fig. 4.8). The velocity reductions are up to about 3% (mostly on v_s) and the minimum values of v_p and v_s in equation 4.4 have been decreased of 0.5% and 0.4%, respectively (see Tab. 4.2). A few small blue features, corresponding to a slight velocity increase ($<$ 0.4%), are also evident both on the surface and inside the simulated volume. However, one should be aware in interpreting these structures at first iterations; if they are reliable, their robustness will be proved by the successive updates of the model.

To have a more clear view of velocity changes between $\mathbf{m}^1$ and $\mathbf{m}^0$, Figure 4.10 shows for v_p and v_s variations (eq. 4.42) a cross-section along the strike of L'Aquila mainshock fault (section 3.2), i.e., a NW-SE section (on the left), and a cross-section across the strike of L'Aquila mainshock fault, i.e., a NE-SW section (on the right). It is well evident that the new model $\mathbf{m}^1$ features velocity decrease with respect to $\mathbf{m}^0$ also in depth. In particular, shear wave variations extend deeper than v_p variations that are mainly concentrated near the surface. This is due to the higher sensitivity in depth of the misfit kernel K_β with respect to K_α, as highlighted in Figure 4.8.

A quantitative assessment of the behaviour of the new model can be obtained, as previously discussed (section 4.6.2), by comparing the total misfit function 4.26

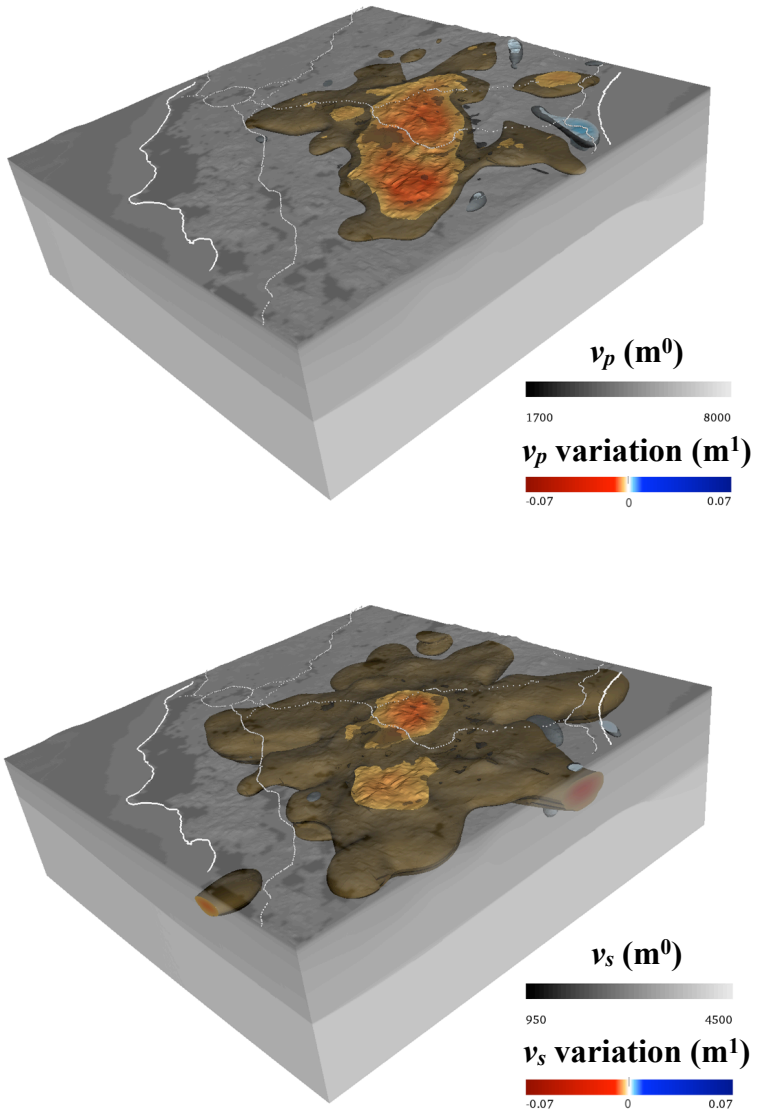


Figure 4.9: Velocity variations for model $\mathbf{m}^1$ - The velocity variations of model $\mathbf{m}^1$ with respect to the initial model $\mathbf{m}^0$ (eq. 4.42) are shown for v_p (top) and v_s (bottom) based upon the color scales. The scale of grey represents the velocities (in m/s) in the initial model $\mathbf{m}^0$ for comparison. The simulated volume in central Italy is viewed from SE and the coastlines and highways are white.

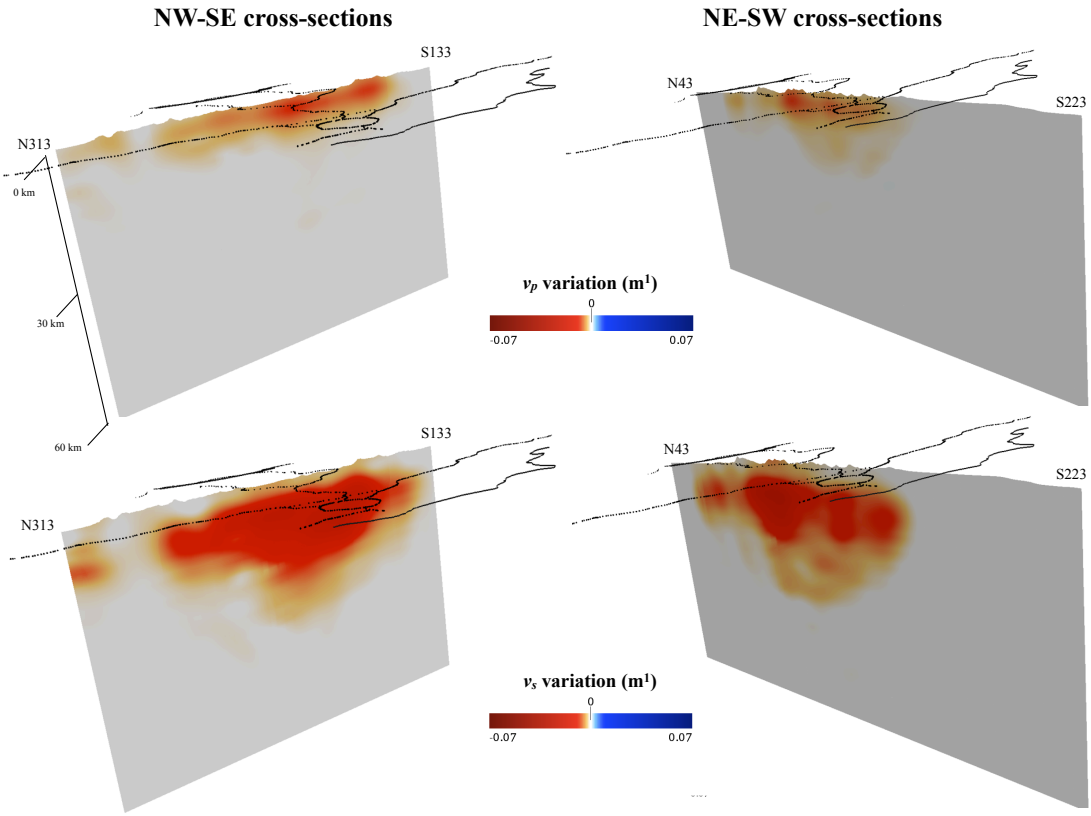


Figure 4.10: Cross-sections of velocity variations for model $\mathbf{m}^1$ - The velocity variations of model $\mathbf{m}^1$ with respect to $\mathbf{m}^0$ (eq. 4.42) are shown for v_p (top) and v_s (bottom), for a section along the L'Aquila event strike (left) and a section across the strike (right). The coastlines and highways are black.

of model $\mathbf{m}^1$ with that of model $\mathbf{m}^0$. The synthetic seismograms associated to the new model are in general different (although slightly) from the synthetics of model $\mathbf{m}^0$ (e.g., Figs. 4.26 - 4.29), and they seek to better fit the data. As a consequence, the time windows selected on the seismograms and the associated measurements could be changed. Thus, we run `FLEXWIN` for $\mathbf{m}^1$ in order to select the new windows, and within them we estimate the new measurements using `measure_adj`. This gives us the estimate of the overall misfit function for $\mathbf{m}^1$, $F(\mathbf{m}^1)$, summed over all the measurement windows, the considered events and the involved categories. It is worth noting that, in order to reduce the computational costs (Fig. 4.38), for the first iterations we focused only on the 11 representative events used for the steplength test (Fig. 4.7 (a)). Thus, to calculate the total misfit $F(\mathbf{m}^1)$ we sum over $N_s = 11$ events; instead, the number of categories is always 6, as explained in section 4.6.2. The results are presented in Figure 4.11 (a), that shows the comparison between the total misfit function for $\mathbf{m}^0$ and that for $\mathbf{m}^1$, both calculated considering only 11 events. One notes that, for this first iteration, the minimization procedure is properly working, since the model update produced a decreasing of the corresponding misfit function.

Some examples of the changes in waveforms for model $\mathbf{m}^1$ with respect to $\mathbf{m}^0$ can be observed in Figures 4.26 - 4.29 (considering only the first two rows of seismograms, labeled by m00 and m01). At this first iteration the variations are very slight. Nevertheless, analysing the whole dataset, one can generally note a decreasing of time shifts between phase arrivals. In particular, in agreement with the velocity reductions shown in Figures 4.9 - 4.10, mostly the positive values of the traveltime differences tend to be decreased, moving the too early synthetic picks toward the data. For the selected measurement windows, the variation of measured time shifts is quantified looking at the boxes above each seismogram component in Figures 4.26 - 4.29. The bars in the boxes indicate, for each model iteration (from $\mathbf{m}^0$ to $\mathbf{m}^5$), the time windows selected in the seismograms, and their colors define the values of the traveltime difference in the window. Based on the color scale in the picture, when color approaches white, the time shift approaches zero and the fits improve. This is evident, e.g., in Figure 4.29, where the bars become white (for T and Z components) updating the model from $\mathbf{m}^0$ to $\mathbf{m}^1$.

The trend of time shift variations for all the seismograms of the 11 events is represented in Figure 4.11 (b), that shows the traveltime measurement distribution for $\mathbf{m}^1$ (red histogram) in comparison with that for $\mathbf{m}^0$ (black contour histogram). As

anticipated, mainly the positive measurements start to be reduced and the distribution begins to move toward a Gaussian centered in zero (the ideal case would be a delta function in zero).

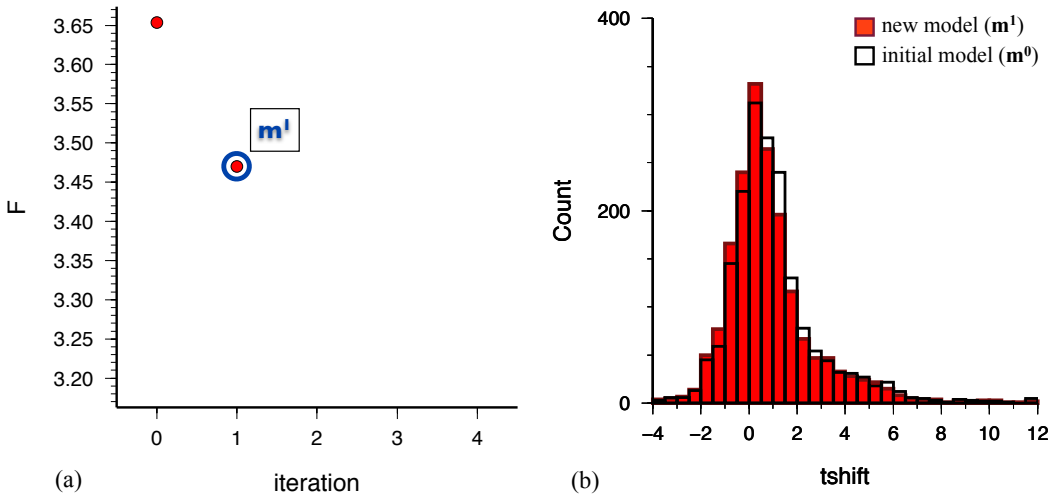


Figure 4.11: Misfit analysis for model $\mathbf{m}^1$ - (a) Traveltime misfit function values (eq. 4.26) for model $\mathbf{m}^1$ with respect to model $\mathbf{m}^0$ (the first point in the graph). Only the 11 events in Figure 4.7 (a) have been considered. (b) Distribution of the traveltime differences (in s) between data and synthetics for model $\mathbf{m}^1$ (red histogram), compared to the distribution for model $\mathbf{m}^0$ (black contour histogram).

4.9 Next iterations and last tomographic model $\mathbf{m}^5$

As shown by the seismogram behaviour for model $\mathbf{m}^1$ (e.g., Figures 4.26 - 4.29), more iterations of the inversion procedure need to be performed in order to reduce the misfit between data and synthetics. To this end, the model at the first iteration $\mathbf{m}^1$ is considered as the new starting model and the steps of the procedure described above are repeated in order to obtain model $\mathbf{m}^2$, and so on, continuing iterating. In general, based upon the steepest descent algorithm, to obtain the improved model at iteration n, $\mathbf{m}^n$, one should update the model at the previous iteration, $\mathbf{m}^{n-1}$, considered as the

starting model, using equation (section 2.3.4.1)

$$\mathbf{m}^n \,=\, \mathbf{m}^{n-1} \,+\, \alpha\, \mathbf{p}^{n-1}\,. \tag{4.43}$$

The steepest descent vector $\mathbf{p}^{n-1}$ at the previous iteration is equal to minus the gradient $\mathbf{g}^{n-1}$ and, thus, it can be obtained from the misfit kernels at iteration $n-1$ through equation (in discrete form; cfr. eq. 4.34)

$$p_k^{n-1} \,=\, -\, g_k^{n-1} \,=\, -\, K_k^{n-1}\, \sqrt{V_k}\,. \tag{4.44}$$

The misfit kernels in 4.44 used for all the next iterations (sections 4.9.1 - 4.9.2) are K_α^{PR-SM} and K_β^{PR-SM} in 4.39, i.e., they have been regularized using both the preconditioning and the smoothing[1], based upon the results in section 4.7.3. Moreover, to obtain the updated models $\mathbf{m}^2$ - $\mathbf{m}^4$ (section 4.9.1) we considered only the 11 representative events, thus the misfit kernels are the sum of the corresponding 11 event kernels (eqs. 4.33 with $N_s = 11$). For the last performed iteration $\mathbf{m}^5$ (section 4.9.2), we tried to reintroduce all the 63 events, thus the kernels of $\mathbf{m}^4$ (the previous iteration) are summed over $N_s = 63$ earthquakes.

To obtain each new iteration, based upon 4.43, every time one needs to calculate the misfit function and its gradient (then $\mathbf{p}$) at the previous iteration, and a steplength test should be performed based upon 4.41 to estimate the suitable α. Note that this test, for all iterations, will always involve only the 11 events.

4.9.1 Iterations $\mathbf{m}^2$ - $\mathbf{m}^4$

Following what has been done for the previous iteration model $\mathbf{m}^1$, in order to obtain the updated model $\mathbf{m}^2$, the adjoint sources for $\mathbf{m}^1$ are calculated using `measure_adj`. Then, SPECFEM3D allows us to construct the event kernels for the 11 events now considered, and, from them, the preconditioned-smoothed misfit kernels for $\mathbf{m}^1$. The steplength test using $\mathbf{p}^1$ and the 11 earthquakes is performed, and a new value of α has been estimated, that allows us to determine the updated model vector $\mathbf{m}^2$. The same procedure has been followed to obtain also $\mathbf{m}^3$ and $\mathbf{m}^4$, except for the fact that in the steplength search for these two cases the tested values of α were (0.02, 0.03, 0.04, 0.05). Thus, we tried to look at the behaviour of the misfit for stronger update percentages,

[1] For all the iterations we used a Gaussian smoothing with $\sigma_x = \sigma_y = 5$ km and $\sigma_z = 1$ km, see section 4.7.2.

but in general the misfit functions tend to re-increase for steplength of 4-5%, justifying the range of tested values.

The choice of considering only 11 events for these iterations is motivated by the need to perform several model updates with a relatively low computational cost, as introduced in section 4.8 and evident in Figure 4.38. Moreover, the 11 earthquakes have been selected (section 4.7.3) so that they were as much representative of the whole dataset as possible, thus it seems reasonable to use them at least for few iterations.

Velocity variations

To observe how the tomographic inversion for central Italy proceeds, one can firstly analyse the velocity variations with respect to the initial model $\mathbf{m}^0$, that are represented in Figures 4.12 - 4.17 for the updated models $\mathbf{m}^2$, $\mathbf{m}^3$ and $\mathbf{m}^4$. The figures show also the variations for $\mathbf{m}^1$ (already presented in section 4.8) and $\mathbf{m}^5$ (discussed in details in section 4.9.2), to have a complete view of the inversion path. In general, the variable represented in Figures 4.12 - 4.17 is

$$\frac{v(\mathbf{m}^n) - v(\mathbf{m}^0)}{v(\mathbf{m}^0)} \,. \tag{4.45}$$

Starting from v_p (Figs. 4.12 - 4.14), the variations extend both on the surface and in depth, and their absolute value continues increasing at each iteration. In particular, the negative-red variations are always dominant and reach values up to about 3% for $\mathbf{m}^4$, indicating that the model update continues decreasing the initial compressional wavespeed almost everywhere. These velocity reductions are mainly concentrated in the central part of the simulated volume, corresponding to Abruzzi and Lazio regions, where most of the sources are located. Moreover, the cross-sections (Figs. 4.13 - 4.14) show that v_p values are decreased mostly in the first 10 km depth, indicating that the sensitivity of kernels K_α used for these iterations is limited to a shallow layer below the surface (as for K_α of $\mathbf{m}^0$, see Fig. 4.8 and section 4.8). Regarding positive v_p variations (increments), most of the blue features evident in $\mathbf{m}^1$ (Figs. 4.9 and 4.12 top left) persist in $\mathbf{m}^2$ - $\mathbf{m}^4$ (Fig. 4.12). Moreover, starting from $\mathbf{m}^2$, a blue structure delineates along the Tevere valley (NW of the simulated domain). In general, v_p increments are lower than reductions and they reach a maximum of almost 1% for $\mathbf{m}^4$.

Considering the shear wavespeed v_s (Figs. 4.15 - 4.17), also for this variable the variations continuously increase their extension and absolute value from $\mathbf{m}^1$ until $\mathbf{m}^4$ (and also $\mathbf{m}^5$, section 4.9.2). Velocity reductions (red features) interest most regions of

the considered volume and they reach a maximum of almost 8% (in $\mathbf{m}^4$) with respect to the initial model $\mathbf{m}^0$. Thus, the decrease percentage of v_s in the inversion is stronger than for v_p and this will affect v_p/v_s ratio, as discussed for the last model in section 4.9.2. Looking at Figures 4.16 - 4.17, one notes that shear wavespeed decrease has a large extension in depth, where the highest percentages are reached at all iterations. In particular, both cross-sections show negative variation patterns that extend down to about 30 km depth in the central part of the volume. On the other hand, increments of v_s with respect to $\mathbf{m}^0$ are relatively low and they are comparable in percentage to v_p increments, reaching about 1% in $\mathbf{m}^4$. The blue feature along the Tevere valley is again evident starting from $\mathbf{m}^2$, but it is less marked than for v_p (cfr. Figs. 4.12 and 4.15). A pronounced blue-positive structure appears along the NW-SE section in Figure 4.16 starting from model $\mathbf{m}^2$. This shallow feature, slightly evident on the surface (Fig. 4.15), corresponds to the Montereale area, NW of L'Aquila.

The minimum and the maximum values of v_p and v_s for the iterations considered in this paragraph are reported in Table 4.2, together with the maximum positive and negative velocity variation percentages. It is worth noting that v_p^{min} and v_s^{min} keep decreasing for successive iterations. Moreover, in general, velocity reductions are always stronger than increments, and they are higher for v_s than for v_p.

Misfit analysis

To quantify the model improvements, one can consider the results of the misfit analysis for iterations $\mathbf{m}^2$ - $\mathbf{m}^4$ (for $\mathbf{m}^5$ see next section 4.9.2).

Figure 4.18 (a) shows the values of the total misfit function 4.26 for models until $\mathbf{m}^4$ (included $\mathbf{m}^1$ discussed in section 4.8) compared to the one for the initial model, $F(\mathbf{m}^0)$. As previously introduced, for these first iterations the misfit functions are obtained by considering only 11 events and are summed over two period bands, 2-20 s and 6-20 s, and three ground motion components. The result of this analysis is that every iteration continues reducing the initial misfit, thus the tomographic inversion tends to improve the starting velocity model, as expected. In particular, a considerable decrease of F results for models until $\mathbf{m}^3$, whereas iteration $\mathbf{m}^4$ shows a more slight reduction.

It is worth noting that the almost stable behaviour of the misfit function at iteration $\mathbf{m}^4$ may suggest that additional information to constrain the inversion is necessary, in order to further reduce the misfit. One possibility is to consider the remaining events together with the other eleven, and thus iteration $\mathbf{m}^5$, discussed in the next section,

will involve all the 63 events. Another possibility is to use the last obtained improved structure model in order to update the initial source parameters. Tape et al. (189) (2010) performed such an inversion at the initial model $\mathbf{m}^0$ and at iteration $\mathbf{m}^{12}$ in the adjoint tomography of California, showing that it introduces remarkable improvements. However, as already mentioned, this is beyond the scope of this work, principally due to its computational costs, and it will be considered for future studies.

Looking at Figure 4.18, one can also observe the behaviour of the traveltime measurement distribution (subfigures (b)-(e)). Starting from $\mathbf{m}^1$ (discussed in section 4.8), the histograms keep tending to Gaussian distributions centered in zero, and mostly the positive traveltime differences decrease, since the velocity values have been mainly reduced in the inversion (Figs. 4.12 - 4.17). This is evident also analysing waveform comparisons for successive iterations represented in Figures 4.26 - 4.29 (see also Figs. 4.30 - 4.37 for model $\mathbf{m}^4$). Firstly looking at the color bars that summarize measurement window behaviour for the models discussed in this section ($\mathbf{m}^2$ - $\mathbf{m}^4$), in most of the examples, increasing iterations, the time shift between phases tend to zero (i.e., the window tend to whitish colors), especially the positive (blue) values. Consistently, focusing on waveforms, one notes that the synthetic seismograms in many cases tend to fit the data more and more accurately in both period ranges (2-20 s and 6-20 s). This means that the improved models are progressively able to better reproduce the features in the real structure that influence seismic wave propagation. However, during the inversion procedure, the fit of some phases may worsen (e.g., picks after about 25 s on T and Z components in Fig. 4.26 and on T component in Fig. 4.33, or pick at about 32 s on Z component in Fig. 4.35) and many features in the seismograms are still not modeled. This should be expected in an inversion that involves different types of waves for different periods and components, and it highlights the need to continue iterating. Moreover, the results reveal the limits of our inversion, since amplitude tomography and source inversion become more and more necessary to improve amplitudes and waveform shapes, slightly adjusted by traveltime tomography.

4.9.2 Last tomographic model $\mathbf{m}^5$

The last updated velocity model that we obtained in this work is $\mathbf{m}^5$. In order to perform this last iteration, based upon the results for previous models (section 4.9.1), we decided to reintroduce all the considered 63 events. Although most of them are

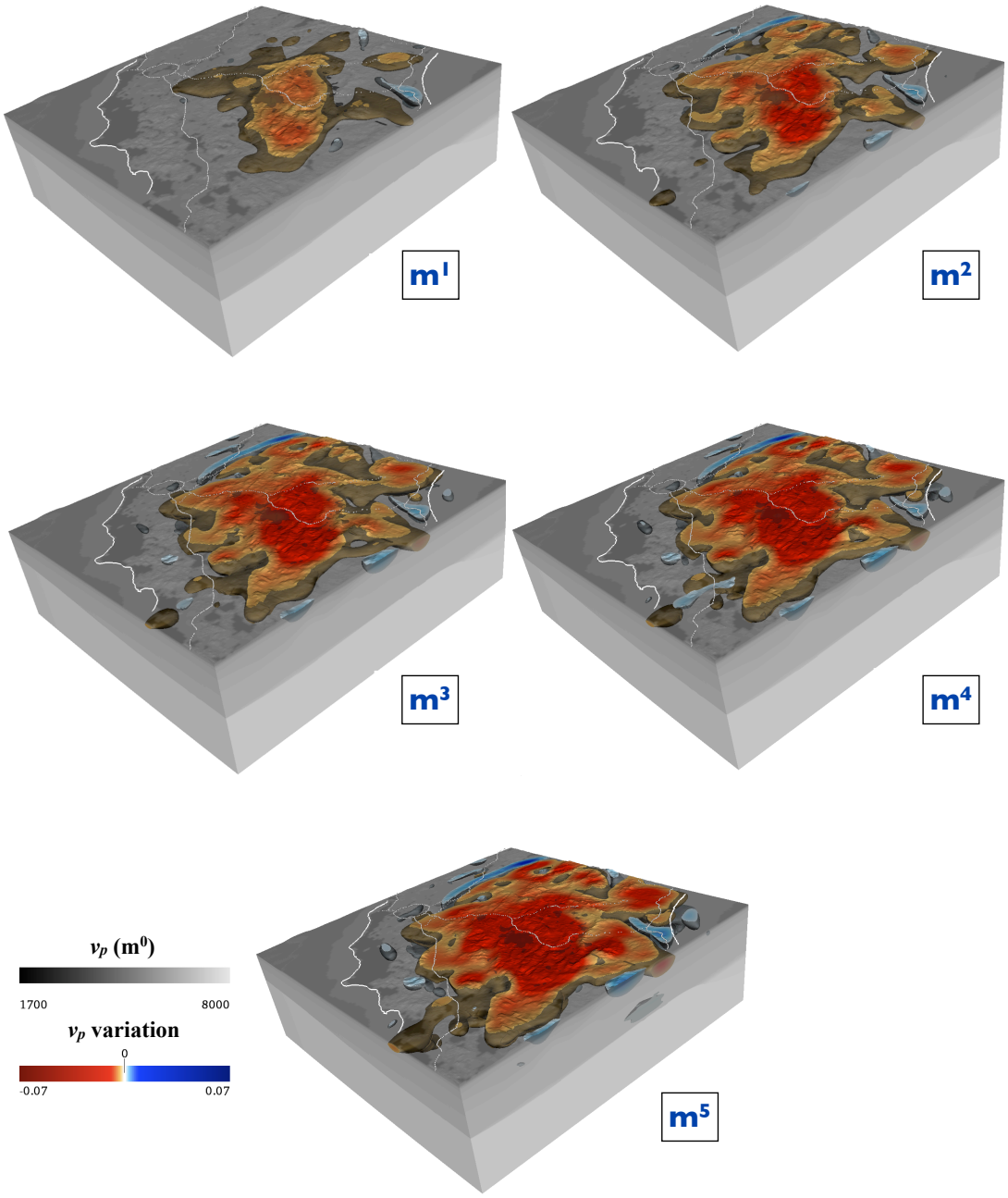


Figure 4.12: v_p variations for the updated models - The compressional wavespeed variations with respect to the initial model $\mathbf{m}^0$ (eq. 4.45) are shown for the five updated models $\mathbf{m}^1$-$\mathbf{m}^5$. See Figure 4.9 for more details.

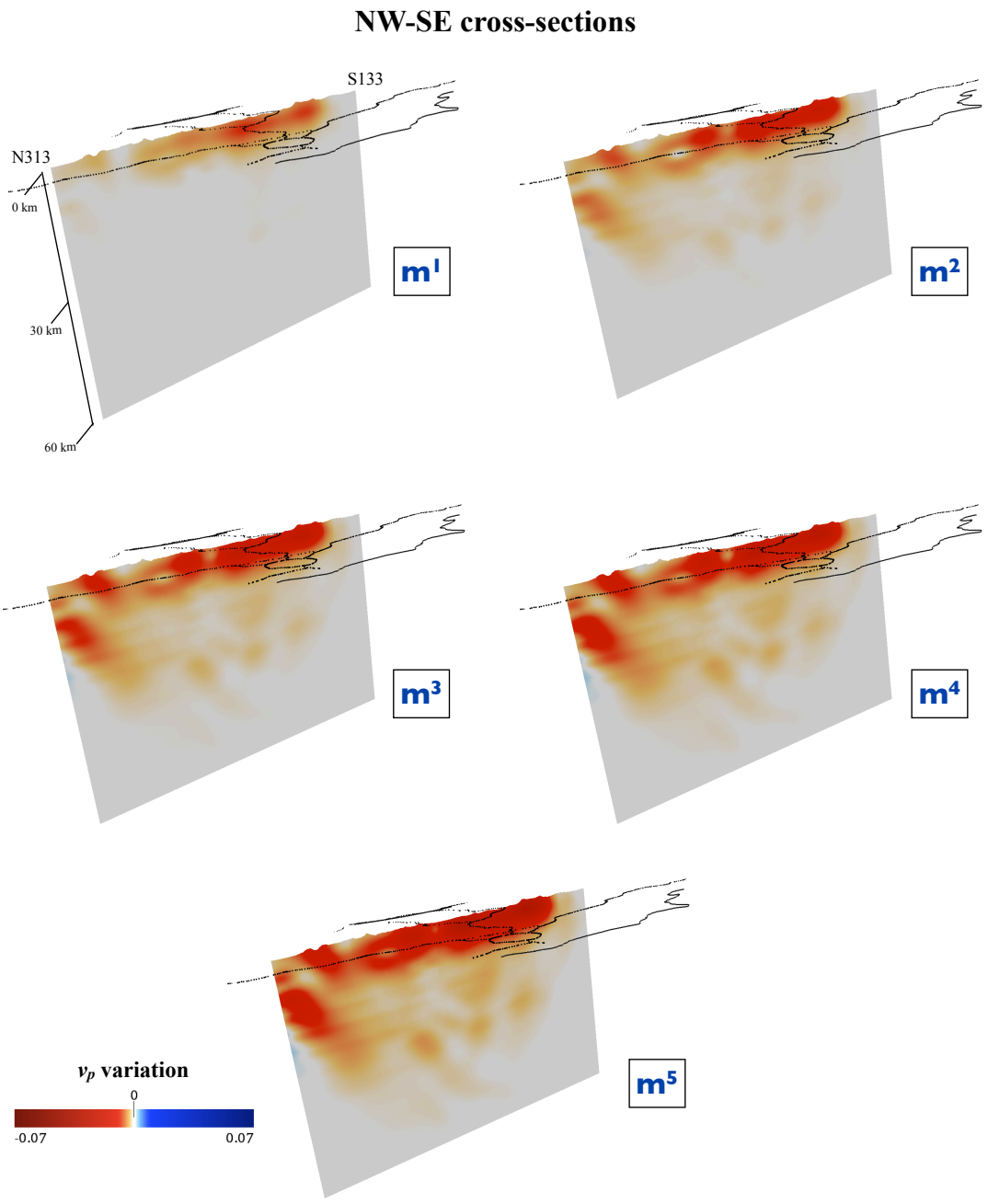


Figure 4.13: Cross-sections of v_p variations for the updated models - The compressional wavespeed variations with respect to the initial model $\mathbf{m}^0$ (eq. 4.45) are shown for the five updated models $\mathbf{m}^1$-$\mathbf{m}^5$ and for a section along the L'Aquila event strike (section 3.2). The coastlines and highways are black.

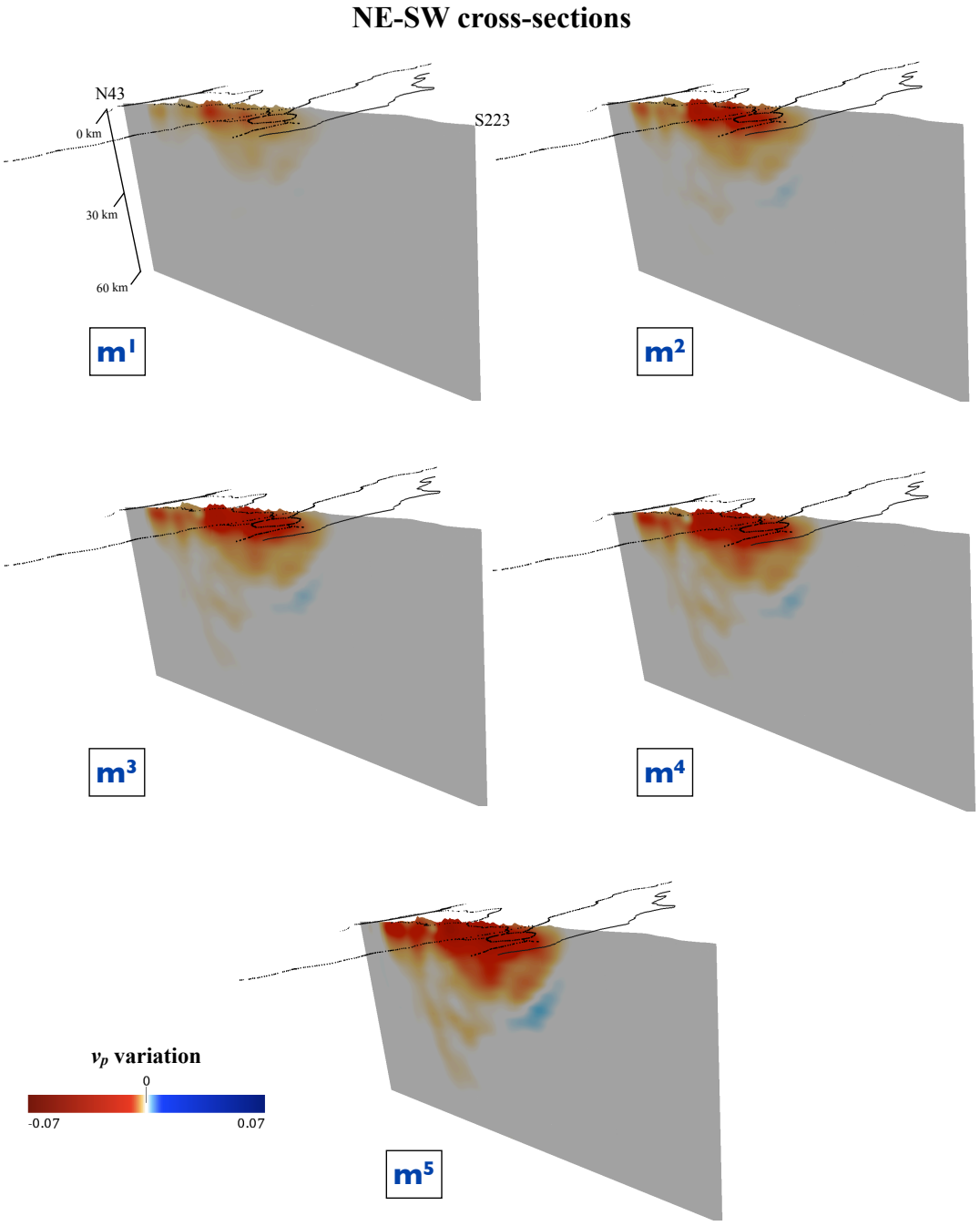


Figure 4.14: Cross-sections of v_p variations for the updated models - Same as Figure 4.13, but for a section across the L'Aquila event strike (section 3.2).

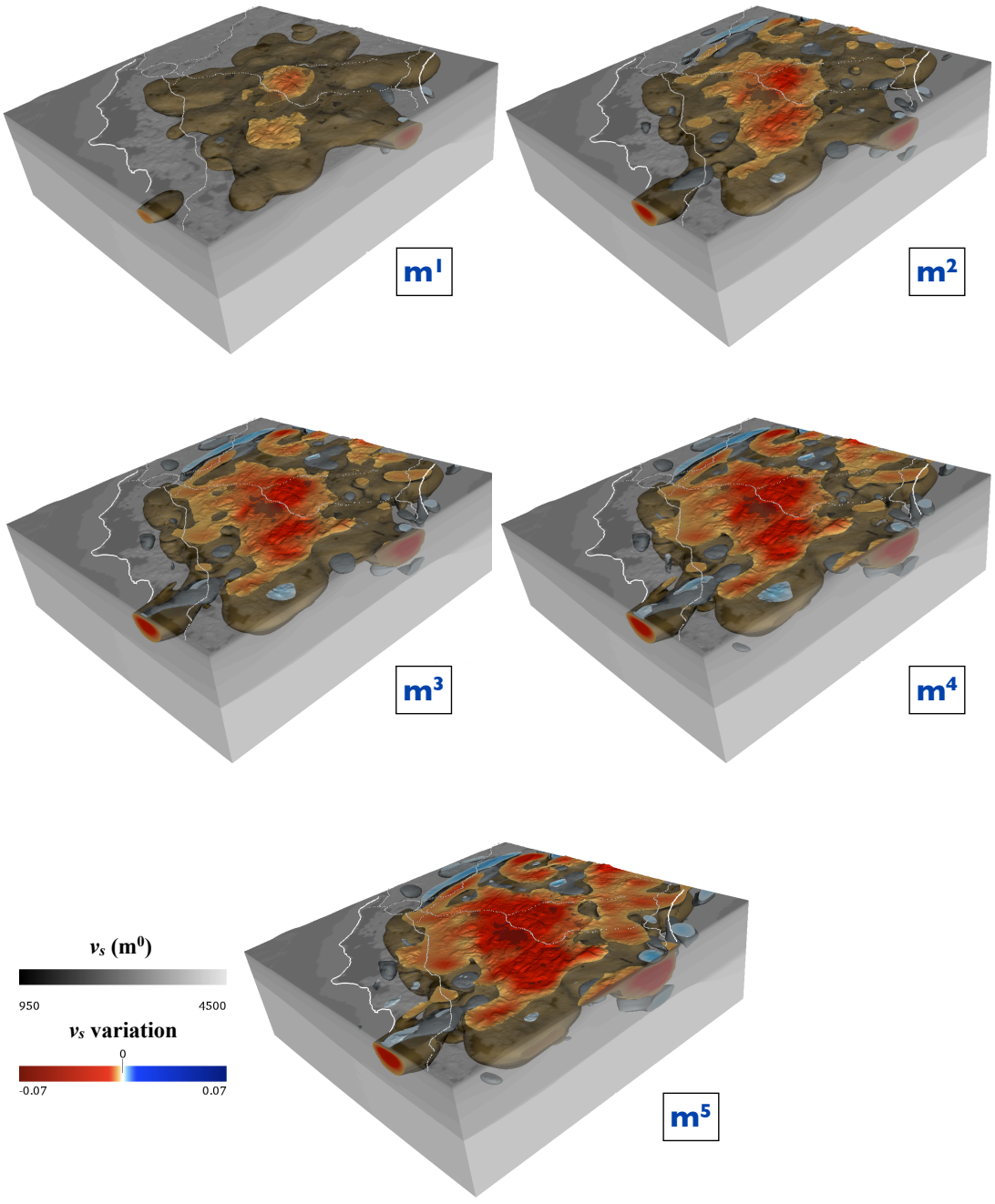


Figure 4.15: v_s variations for the updated models - Same as Figure 4.12, but for the shear wavespeed v_s.

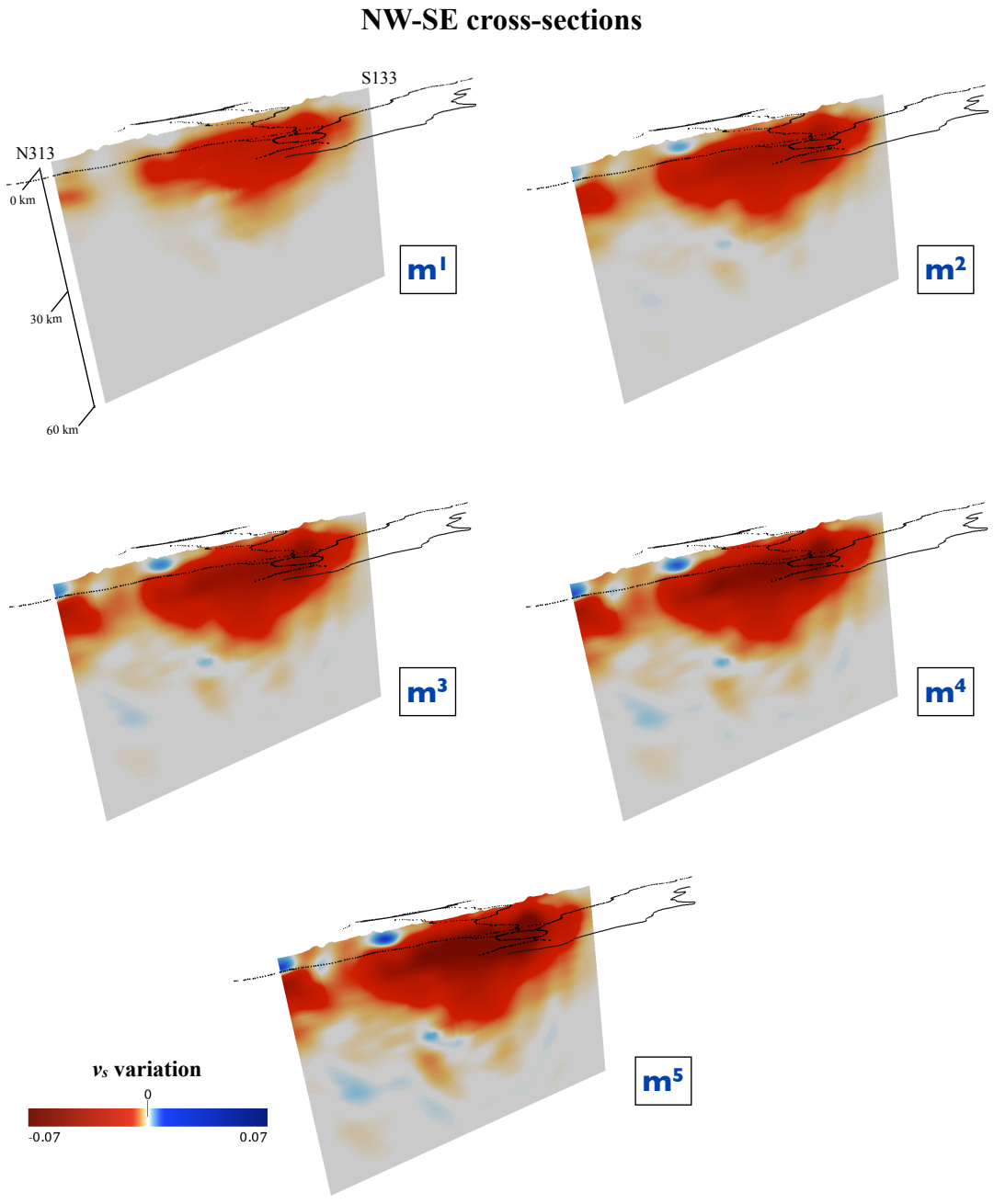


Figure 4.16: Cross-sections of v_s variations for the updated models - Same as Figure 4.13, but for the shear wavespeed v_s.

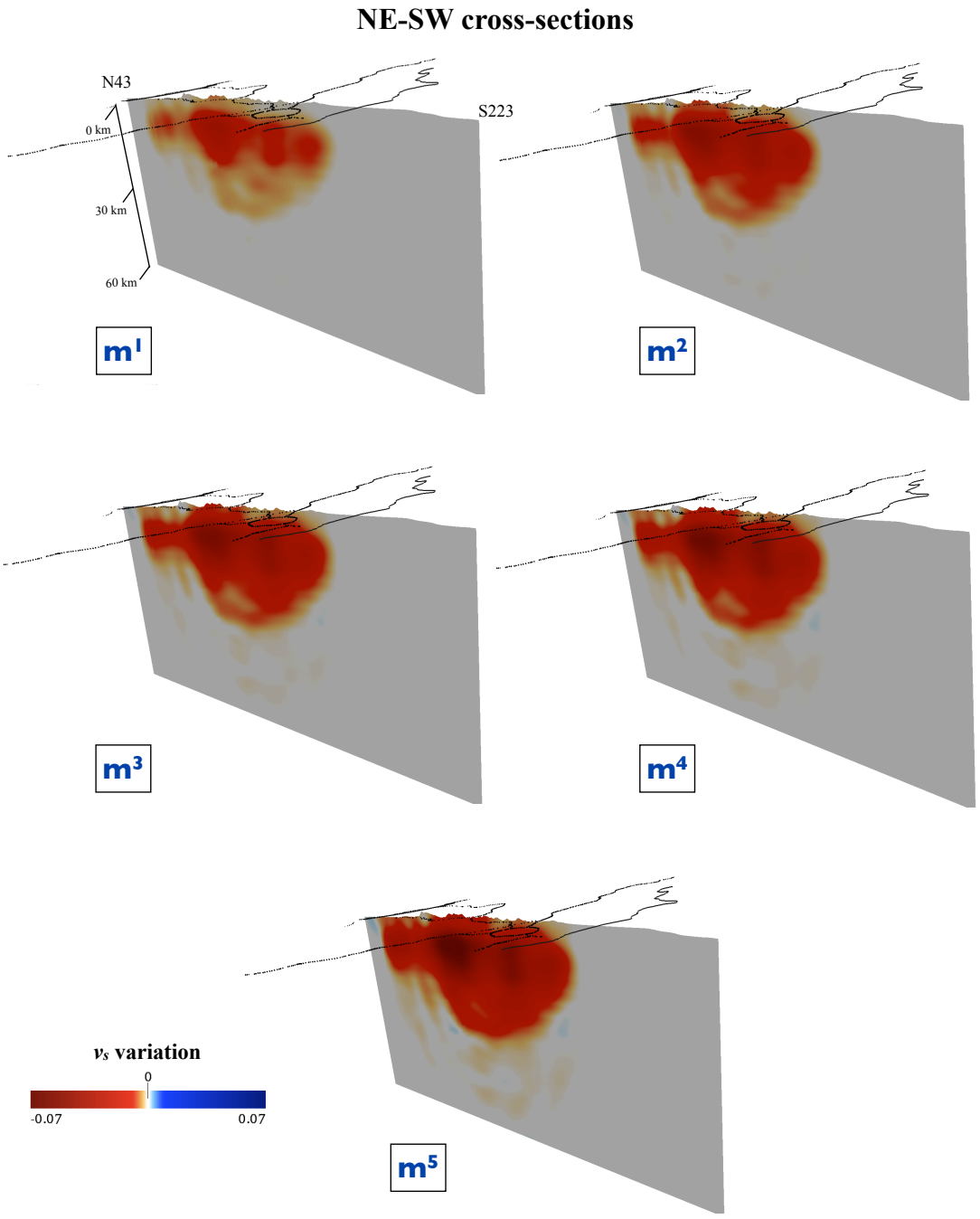


Figure 4.17: Cross-sections of v_s variations for the updated models - Same as Figure 4.14, but for the shear wavespeed v_s.

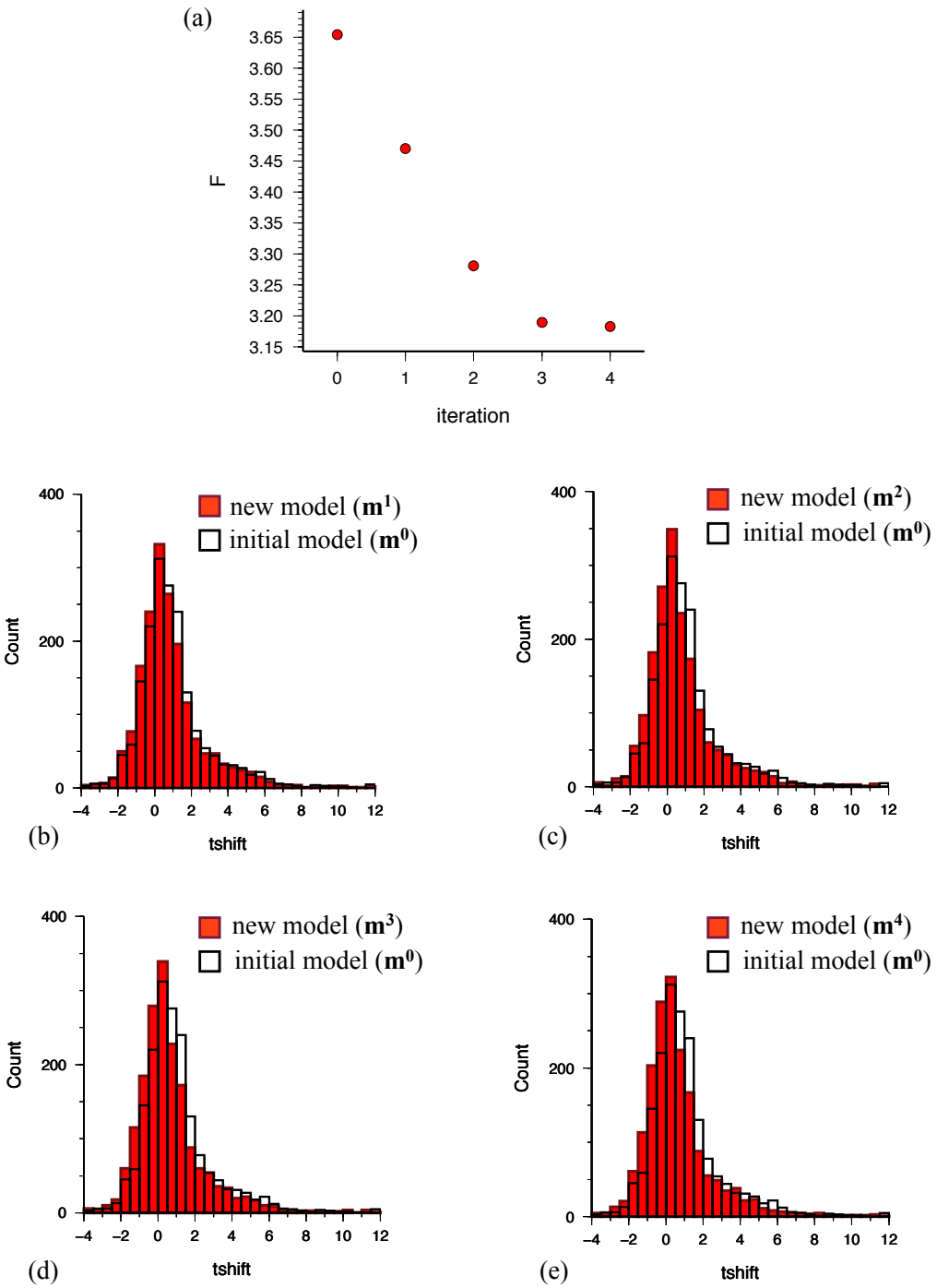


Figure 4.18: Misfit analysis for the updated models $\mathbf{m}^1$-$\mathbf{m}^4$ - (a) Traveltime misfit function values (eq. 4.26) for the updated models $\mathbf{m}^1$-$\mathbf{m}^4$ with respect to model $\mathbf{m}^0$ (the first point in the graph). Only the 11 events in Figure 4.7 (a) have been considered. (b)-(e) Distribution of the traveltime differences (in s) for model $\mathbf{m}^1$ (red histogram fig.(b)), $\mathbf{m}^2$ (red histogram fig.(c)), $\mathbf{m}^3$ (red histogram fig.(d)) and $\mathbf{m}^4$ (red histogram fig.(e)), compared to the distribution for model $\mathbf{m}^0$ (black contour histogram in figs. (b)-(e)).

	$\mathbf{m}^0$	$\mathbf{m}^1$	$\mathbf{m}^2$	$\mathbf{m}^3$	$\mathbf{m}^4$	$\mathbf{m}^5$
v_p^{min} [m/s]	1753.3	1744.0	1740.0	1737.6	1732.5	1722.0
v_p^{max} [m/s]	8000.0	8004.0	8006.6	8008.9	8010.0	8013.5
v_p var [%]		-1.2÷0.3	-2.2÷0.6	-2.8÷0.9	-3.3÷1.1	-4.1÷1.2
v_s^{min} [m/s]	984.5	980.7	978.0	976.0	974.0	971.3
v_s^{max} [m/s]	4444.4	4446.8	4448.9	4452.0	4454.5	4455.9
v_s var [%]		-3.0÷0.3	-5.7÷0.6	-7.1÷0.9	-8.3÷1.0	-10.0÷1.1

Table 4.2: Limits and variations of the velocities - The minimum and maximum values of v_p and v_s are indicated for the initial model $\mathbf{m}^0$ and for the updated models $\mathbf{m}^1$-$\mathbf{m}^5$. The range of the velocity variations (v var) for each updated model with respect to the initial one (eq. 4.45) are also reported (in percentage).

clustered around L'Aquila region, considering the whole dataset will extend kernel sensitivity, adding constraints to the inversion, and this could help the minimization procedure (section 4.9.2.2).

As introduced in section 4.9, the event kernels of the 63 events for model $\mathbf{m}^4$ have been calculated, and the misfit kernels resulting from their sum are used to obtain $\mathbf{p}^4$ in equation 4.43 that gives $\mathbf{m}^5$.

It is fundamental to highlight that $\mathbf{m}^5$ is not the definitive model for central Italy; nevertheless, it shows interesting differences with respect to the initial model $\mathbf{m}^0$ and improved data fittings, as discussed in what follows.

4.9.2.1 Velocity variations and patterns of v_p, v_s and v_p/v_s

As for previous iterations, to present the new model, one can firstly outline the main characteristics of velocity variations in $\mathbf{m}^5$ with respect to $\mathbf{m}^0$ (Figs. 4.12 - 4.17).

Concerning v_p (Figs. 4.12 - 4.14), velocity reductions (red areas) extend again their pattern both on the surface and inside the simulated volume, following the trend discussed for previous iterations and reaching a maximum percentage of 4% (Tab. 4.2). In particular, v_p decrease remains confined within about 10 km depth almost everywhere. A prominent red-negative feature, grown progressively from the first iteration $\mathbf{m}^1$, is evident on the surface around the Maiella area, on the SE of the volume (Fig. 4.12). This is also visible in the NW-SE cross-section of Figure 4.13.

Compressional wavespeed increases (blue features) interest only small regions of the simulated volume and also for this model $\mathbf{m}^5$ their percentage is lower than reduction, maintaining a maximum value of about 1% (Tab. 4.2). The feature along the Tevere valley is stable, whereas few small scalelength structures in the domain disappear, suggesting they were spurious or still unresolved.

Shear wavespeed reduction pattern with respect to $\mathbf{m}^0$ (Figs. 4.15 - 4.17) is similar to the previous iteration $\mathbf{m}^4$ (section 4.9.1), but continues extending, and its maximum value increases up to about 10% (Tab. 4.2). As for v_p, the Maiella area is characterized by a remarkable decrease of v_s. Looking at the cross-sections 4.16 - 4.17, it is evident that v_s decrease extends very deep (down to about 30-35 km) in the central part of the volume (where the highest reduction percentages are reached), highlighting once again the stronger sensitivity of K_β misfit kernel with respect to K_α.

Positive-blue variations of v_s are slightly changed with respect to $\mathbf{m}^4$ and their maximum is again 1% compared to $\mathbf{m}^0$ (Tab. 4.2). In particular, one observes the almost stable weak feature along the Tevere valley and the stronger blue structure in correspondence of Montereale area (Figs. 4.15 - 4.16).

The minimum values of v_p and v_s decrease also for model $\mathbf{m}^5$ (compared to $\mathbf{m}^0$), as expected, with reduction percentages of 1.8% and 1.3%, respectively (Tab. 4.2). The maximum values of v_p and v_s, corresponding to the homogeneous layer below the Moho, are slightly increased in $\mathbf{m}^5$ ($\sim 0.2\%$). However, the Moho discontinuity is kept fixed at each iteration and both kernels have a very weak sensitivity at these depths.

Summarizing the results, a common characteristics of both v_p and v_s variations in the inversion is a reduction of the initial velocity in correspondence of most of the mountain reliefs, as Gran Sasso, Maiella, Sirente and Simbruini massifs. This suggests that the initial model overestimated velocities in these areas, although we already constructed $\mathbf{m}^0$ by reducing upper crustal velocities with respect to Chiarabba et al. (39) (2010)'s tomography (see sections 3.4.1 and 4.2.1). An exception is the Montereale Mt, where we highlighted a sensible increase of v_s, and also of v_p although much more weaker.

Another general feature to be noted is the velocity decrease (both v_p and v_s) in the updated model in correspondence of the main sedimentary basins in central Italy, e.g., Fucino, L'Aquila and Sulmona basins. Chiarabba et al. (39) (2010)'s tomography already showed low v_p (and also v_s) velocities in these areas, and we superimposed the

v_s30 layer to their model in order to construct $\mathbf{m}^0$ (sections 3.4.1 and 4.2.1). However, our minimization procedure suggests the need for a further velocity decrease.

Thus, in general, except for few areas, the inversion tends to decrease the initial velocities. A remarkable point of the results is that these reductions for v_p are mainly concentrated in a thin shallow crustal layer, mainly within the first 10 km depth. On the contrary, for v_s, velocity reductions interest a thick layer from the surface down to about 30 km depth. This is due to the very shallower resolution power of kernel K_α with respect to K_β, used at each iteration of the inversion (for model $\mathbf{m}^0$ see Fig. 4.8). Moreover, v_s decrements are in general stronger than those of v_p, in most of the interested areas.

All the above aspects can be well observed also in Figures 4.19 - 4.24, where we show the six cross-sections analysed by Chiarabba et al. (39) (2010; Fig. 6). We use these profiles to present the updated model $\mathbf{m}^5$, since the main changes with respect to the initial model can be highlighted especially in depth. The characteristics of the new model on the surface are better evidenced looking at the velocity variations already discussed (Figs. 4.12 and 4.15), since the change percentages are lower than in depth. The considered sections sample almost the whole tomographic domain and for each of them we present v_p, v_s and v_p/v_s for $\mathbf{m}^5$ compared to $\mathbf{m}^0$.

The main results for v_p, confirming what summarized for its variations, show that the velocity pattern in depth is almost unchanged with respect to the initial model. Thus, in each section the structures at high v_p (6.7-7.0 km/s) between about 8 and 16 km depth, evidenced by Chiarabba et al. (39) (2010), seem to persist beneath the Simbruini Mts and the Maiella-Gran Sasso thrusts. Moreover, Figures 4.22 and 4.23 show the low v_p anomaly near the Fucino basin located by Chiarabba et al. (39) (2010) at about 12 km depth, as well as a low v_p area near Alban Hills volcano persists in Figure 4.23 (Chiarabba et al. (39), 2010). These features have barely perceptible velocity reductions compared to $\mathbf{m}^0$. The major decrements occur, instead, near the surface in correspondence of the basins, where the minimum values of v_p are already featured by $\mathbf{m}^0$ (e.g., Fig. 4.20).

Concerning v_s, on the other hand, velocity reductions are remarkable and all the cross-sections present a relevant variation of the shear wavespeed pattern. In particular, the deep features (8-16 km depth) with high v_s (around 4 km/s) evident in $\mathbf{m}^0$ (and

corresponding to the high v_p features) are strongly reduced in $\mathbf{m}^5$ and the layer with $v_s \approx$ 3.0-3.2 km/s extends largely in depth.

The different behaviour of v_p and v_s variation patterns and percentages, described so far, results in a remarkable change of the v_p/v_s model from $\mathbf{m}^0$ to $\mathbf{m}^5$. In all the cross-sections 4.19 - 4.24, mainly the central part of the domain features an increase of v_p/v_s ratio compared to the initial model, with a maximum percentage of 10%. v_p/v_s increments extend also in depth and they correlate very well with v_s reductions, as expected from the above discussion. In fact, due to the limited resolution of v_p in depth and the stronger variations of v_s compared to v_p, v_p/v_s ratio is mostly affected by v_s decrease. Moreover, to construct the initial velocity model, the v_p/v_s ratio has been fixed on the surface (section 3.4.1); the different update of v_p and v_s models in the inversion, causes this ratio to be no longer constant at iteration $\mathbf{m}^5$.

High values of v_p/v_s (up to a maximum of 2) are mainly in correspondence and beneath (down to 12 km depth) the principal sedimentary basins, as already obtained by Chiarabba et al. (39) (2010)'s tomography, that however features lower values (up to about 1.92). Moreover, e.g. in cross-section 4.23, the low v_p/v_s anomaly at the Alban Hills volcano shown by Chiarabba et al. (39) (2010) is still present, but the values of the ratio are slightly increased.

An interesting feature of model $\mathbf{m}^5$ is that a high v_p/v_s anomaly is located both on the footwall and the hanging wall of the L'Aquila mainshock fault (Fig. 4.22). Chiarabba et al. (39) (2010)'s tomography is based on earthquakes occurred few years before the L'Aquila event and locates the mainshock hypocenter on top of a high v_p/v_s volume. They interpret this structure as a fluid-filled volume. Several recent studies present variations of seismic properties in correspondence of the 6th April L'Aquila event. Zaccarelli et al. (207) (2011), using ambient noise cross-correlation, infer a velocity reduction of 0.3% in the crust, as a result of the mainshock. Moreover, Lucente et al. (120) (2010), analysing v_p/v_s and shear wave splitting behaviour right before the L'Aquila mainshock, suggest a fluid transfer across the fault zone, from the footwall to the hanging wall, just after the 30th March foreshock. This causes a sharp increase of v_p/v_s also in the hanging wall. In our analysis we use both foreshocks and aftershocks of L'Aquila event, thus we are not able to distinguish v_p/v_s behaviour before ad after the mainshock. However, the number of considered aftershocks is larger than that of foreshocks (section 4.3), thus one may suggest that the structural properties after

the L'Aquila event have a dominant influence on the retrieved tomographic model. A detailed analysis before the mainshock is not possible at present, since we have a too low number of earthquakes in the preceding period. Once the resolution of our problem will be increased, we could include smaller magnitude events, also calculating their moment tensors, and the pre-mainshock phase could be more accurately investigated.

As discussed above, iteration $\mathbf{m}^5$ of our tomographic inversion, primarily based on events after the L'Aquila mainshock, shows, in many regions of the model, values of v_p/v_s higher than those in $\mathbf{m}^0$ (Figs. 4.19 - 4.24). Di Luccio et al. (56) (2010), considering events from October 2008 until April 2009, observe an abrupt increase of v_p/v_s from the 30th March foreshock, with values reaching about 1.95 after the mainshock (consistent with fluid-rich zones; Zhao & Negishi (208), 1998, Husen & Kissling (86), 2001). An increase of v_p/v_s is also observed by Lucente et al. (120) (2010), using events from January until April 6th 2009. After March 30th they found an average v_p/v_s in the epicenter area larger than 1.92, mostly caused by v_s decrease. In our results high v_p/v_s ratios, with peaks of 2, are shown, mainly due to v_s reductions with respect to $\mathbf{m}^0$. Based on the above cited studies, these large values of v_p/v_s may be considered plausible, especially if fluids are supposed to be present. However, this results could be affected by the limits of our inversion procedure. P-wave sensitivity, represented by the kernels K_α, is probably dominated by surface waves that see the P-SV coupling. At the considered short periods (2-20 s), the surface waves are more sensible to shallower structures, thus the resolution of compressional wave kernels in our tomography is limited to a thin superficial layer. On the contrary, shear wave kernels K_β have a deeper sensitivity, as already discussed. This results in poorly resolved v_p variations in depth compared to strong v_s reductions, that in turn cause a probably too high increase of v_p/v_s in the corresponding regions.

In conclusion, it is worth noting that, although model $\mathbf{m}^5$ improves misfit reduction and data fitting (section 4.9.2.2), it is still just an iteration and not the final model. Thus, in general, we prefer not to stress on the geologic and tectonic interpretation of the evidenced model features, since, though some of them are stable, more iterations are required to prove their reliability and robustness.

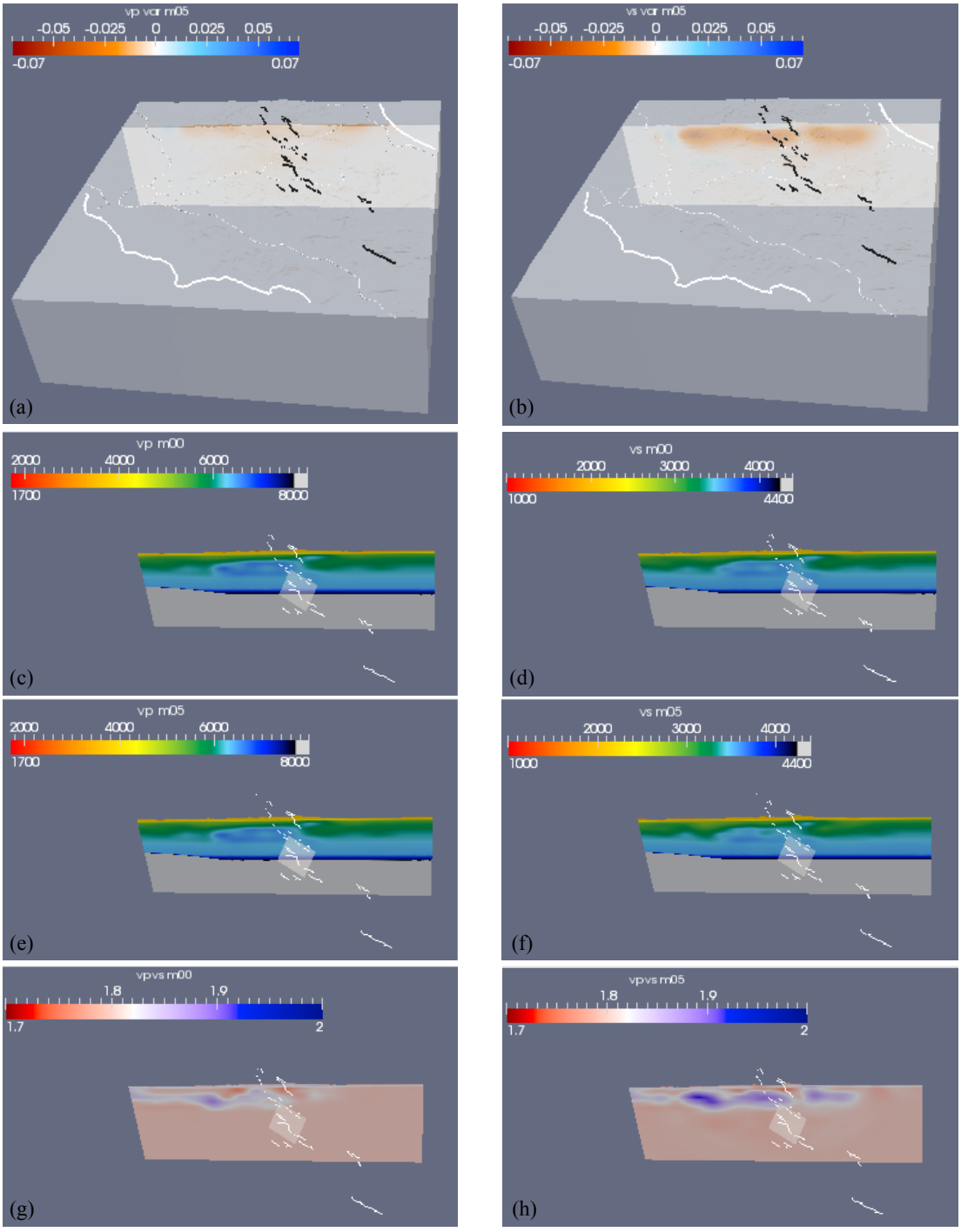


Figure 4.19: Cross-sections for the updated model $\mathbf{m}^5$ - (a)-(b) v_p and v_s variations for model $\mathbf{m}^5$ with respect to $\mathbf{m}^0$ (eq. 4.45) along the indicated cross-section in central Italy. Coastlines and highways are white, principal fault traces are black. (c)-(d) v_p and v_s profiles along the section for model $\mathbf{m}^0$. The principal fault traces are white and the fault plane of the 2009 L'Aquila event is shown in transparency. (e)-(f) v_p and v_s profiles along the section for model $\mathbf{m}^5$. (g)-(h) v_p/v_s ratio for models $\mathbf{m}^0$ and $\mathbf{m}^5$, respectively. The scales of the velocities are in m/s.

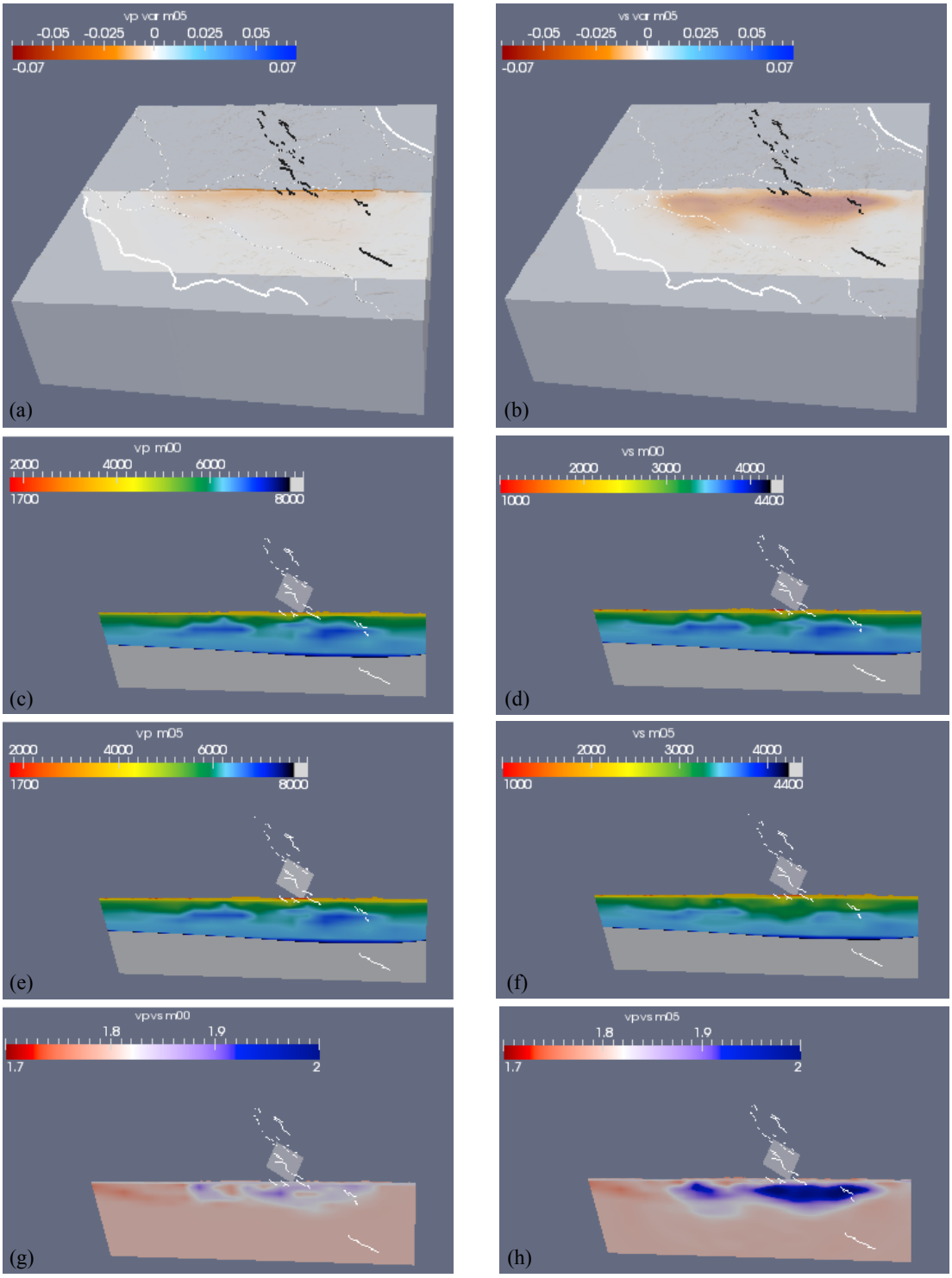


Figure 4.20: Cross-sections for the updated model $\mathbf{m}^5$ - Same as Figure 4.19, but for the cross-section in figs. (a)-(b).

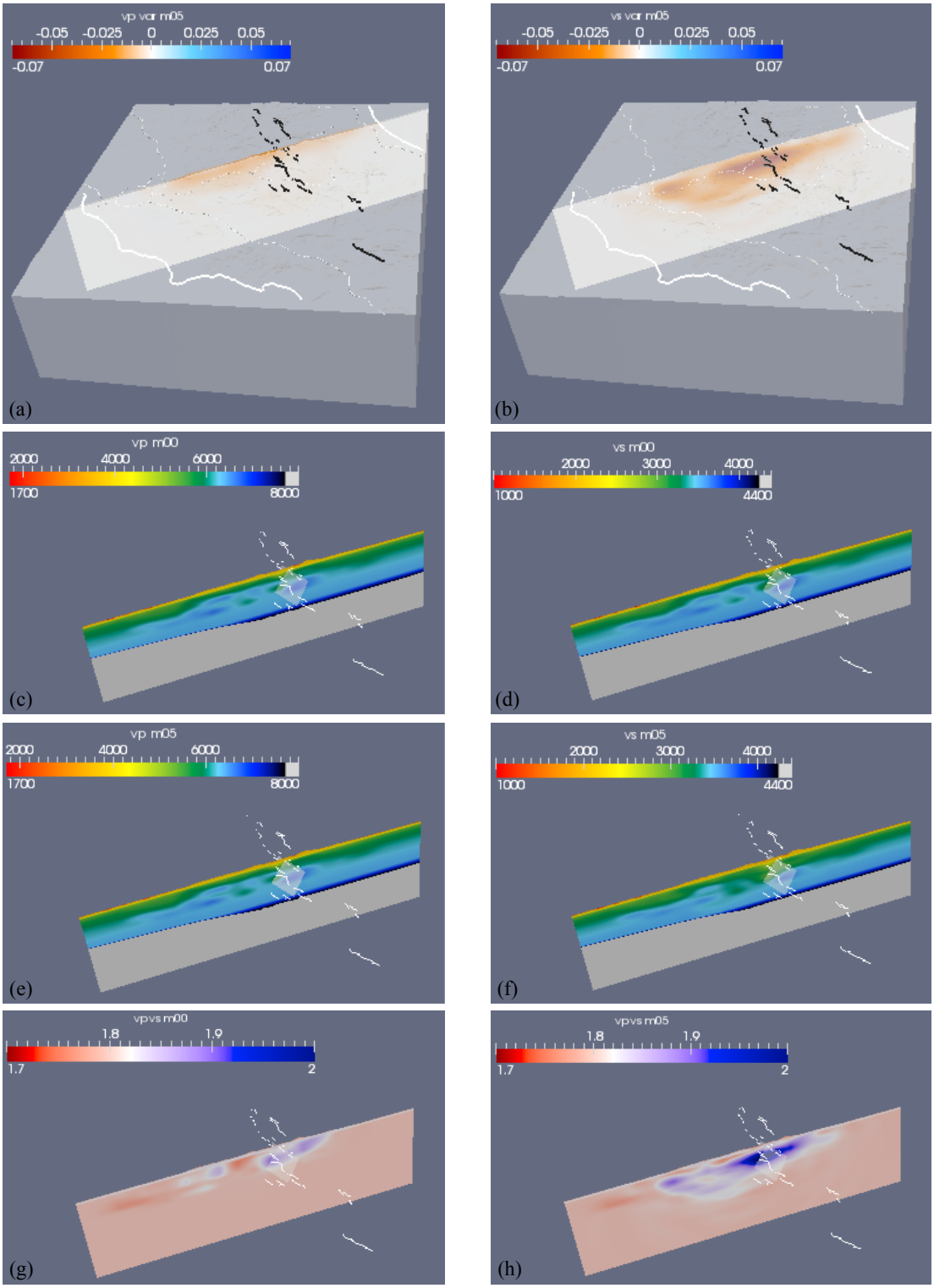


Figure 4.21: Cross-sections for the updated model $\mathbf{m}^5$ - Same as Figure 4.19, but for the cross-section in figs. (a)-(b).

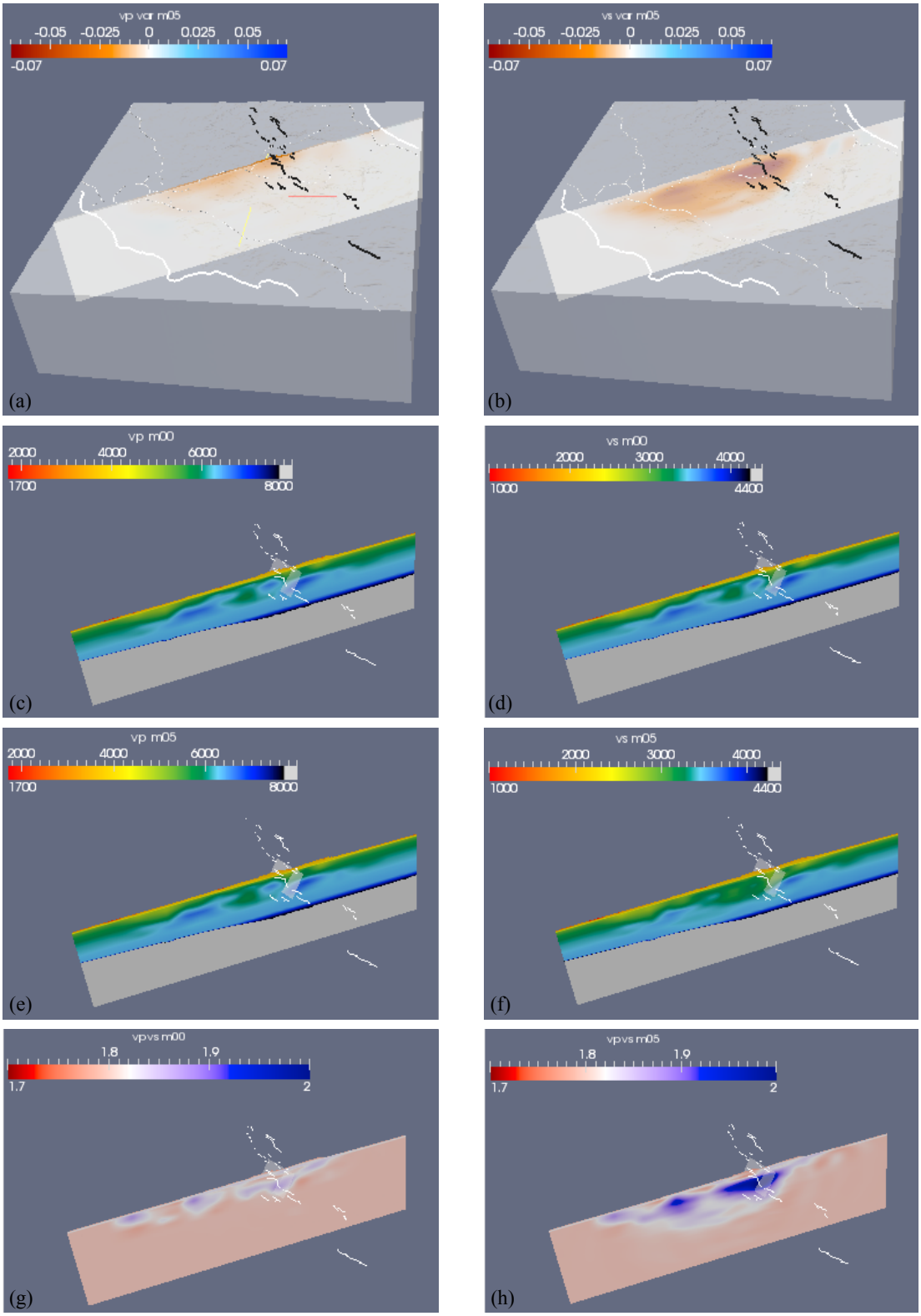


Figure 4.22: Cross-sections for the updated model $\mathbf{m}^5$ - Same as Figure 4.19, but for the cross-section in figs. (a)-(b).

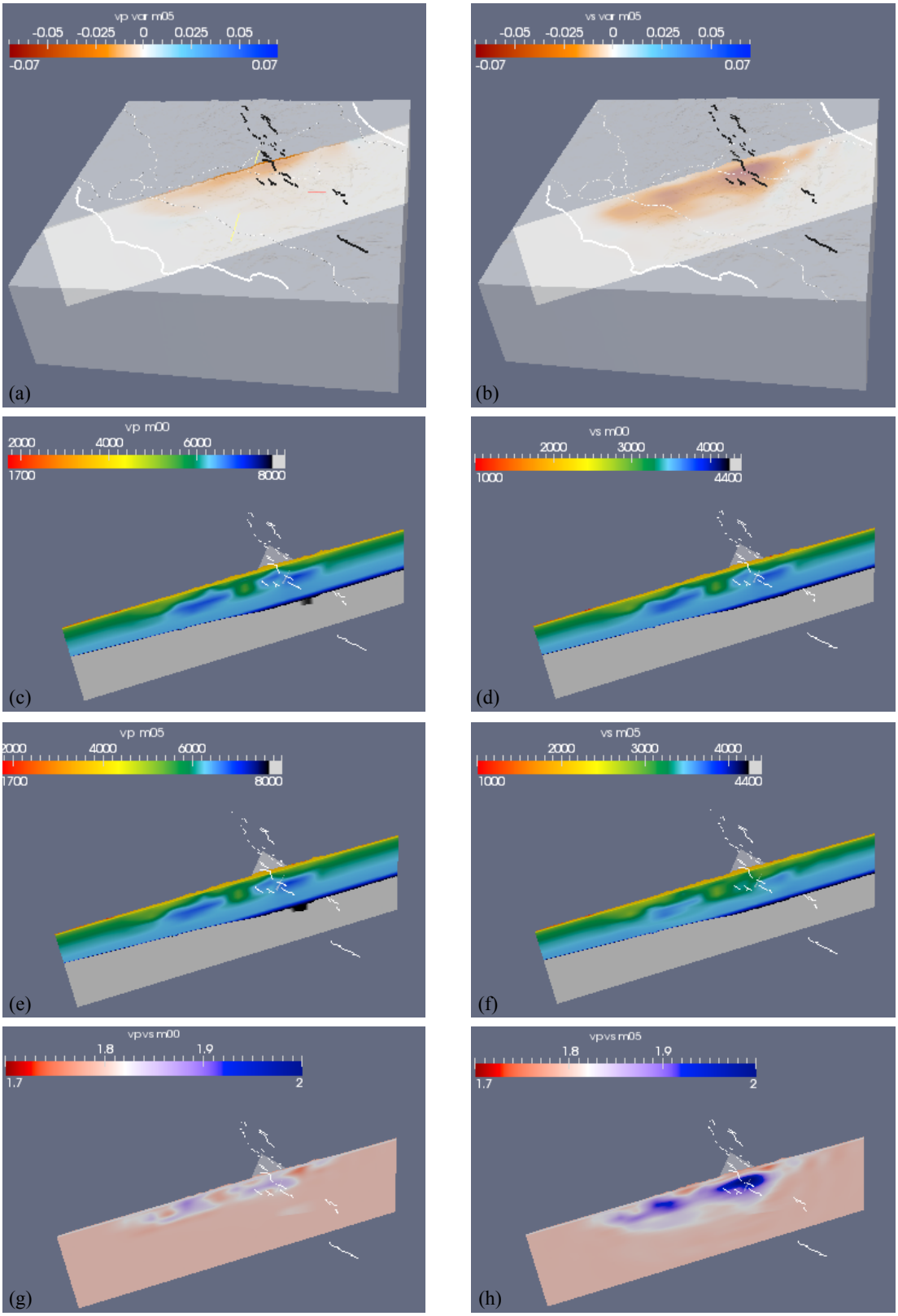


Figure 4.23: Cross-sections for the updated model $\mathbf{m}^5$ - Same as Figure 4.19, but for the cross-section in figs. (a)-(b).

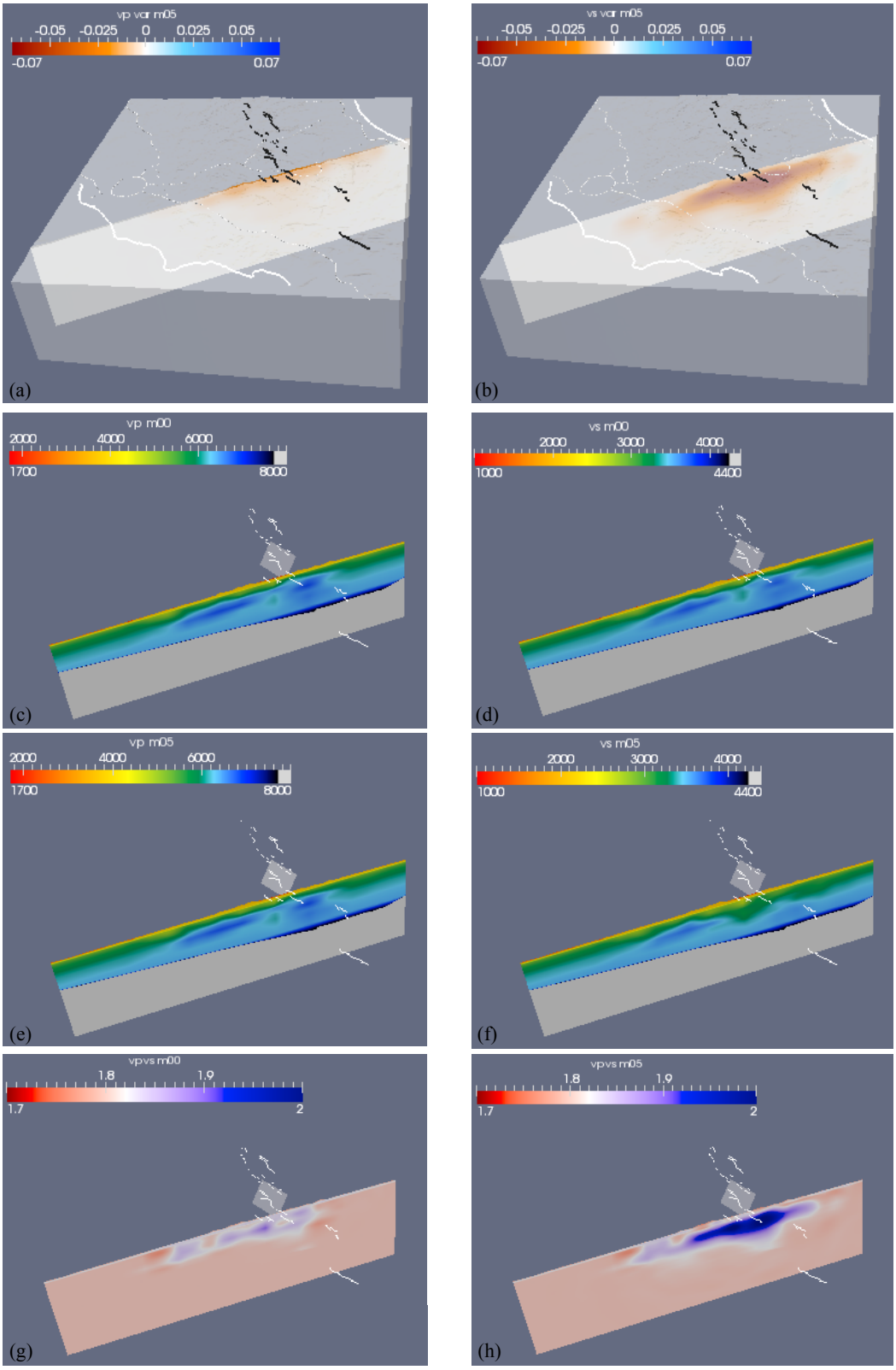


Figure 4.24: Cross-sections for the updated model $\mathbf{m}^5$ - Same as Figure 4.19, but for the cross-section in figs. (a)-(b).

4.9.2.2 Misfit analysis and seismograms

For the last model $\mathbf{m}^5$ the update is obtained by considering all the 63 events, thus also the misfit analysis has been performed based upon the whole dataset. As a consequence, in order to check the minimization trend, we recalculated the total misfit function 4.26 also for $\mathbf{m}^0$ and $\mathbf{m}^4$ by using all the 63 earthquakes. Figure 4.25 (a) shows the comparison between $F(\mathbf{m}^0)$, $F(\mathbf{m}^4)$ and $F(\mathbf{m}^5)$. The result is that the new model $\mathbf{m}^5$ produces a further decrease of the misfit, suggesting that considering all the events in the inversion has introduced valuable information that enhances the convergence.

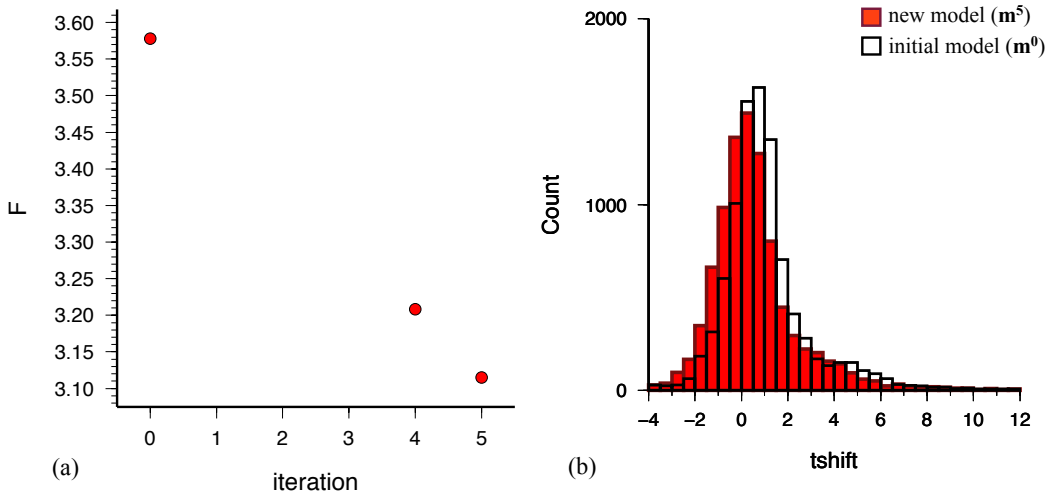


Figure 4.25: Misfit analysis for the whole dataset - (a) Traveltime misfit function values (eq. 4.26) for model $\mathbf{m}^4$ and $\mathbf{m}^5$ with respect to model $\mathbf{m}^0$ (the first point in the graph). All the 63 events in Figure 4.3 (a) have been considered. (b) Distribution of the traveltime differences (in s) between data and synthetics for model $\mathbf{m}^5$ (red histogram), compared to the distribution for the initial model $\mathbf{m}^0$ (black contour histogram).

The total distribution of all traveltime measurements for $\mathbf{m}^5$ with respect to $\mathbf{m}^0$ is shown in Figure 4.25 (b). Its shape, closer to the Gaussian in zero, highlights an evident reduction mostly of positive traveltime differences, following the inversion trend discussed for previous iterations (sections 4.8 - 4.9.1). As a consequence, one expects that mainly the too early synthetic arrivals tend to shift and to improve the fit to the data. Several examples are shown in Figures 4.26 - 4.29, where all the five iterations

are represented for comparison. Synthetic waveforms change slightly from one model to the next one, as previously discussed. However, after five steps of the inversion, the variations with respect to the initial model are evident (see also Figs. 4.30 - 4.37).

In general, one notes that mostly the picks corresponding to the selected measurement windows are adjusted in successive iterations. Nevertheless, these changes may influence the whole seismograms and new phases can be matched (see, e.g., R component for Figs. 4.30 and 4.36, and Z in Fig. 4.33). Hence, the length of the windows and/or their number generally increase updating the model. As a consequence, a larger part of the seismograms is progressively involved in the inversion, aiming to a full waveform fitting.

All the aspects discussed above are evident in most of the Figures 4.26 - 4.37 looking at the boxes with the colored bars that summarize measurement window behaviour. Model $\mathbf{m}^5$, in general, continues updating previous iterations. Thus, for most of the stations and components, while the fitting improves, the windows extend, covering a larger part of the timeseries. Moreover, the bars, based on the color scale, show how the values of traveltime differences for each window change for successive iterations until $\mathbf{m}^5$, pointing out that mostly the positive (blue) differences tend to become zero (white), since mainly velocity reductions result from model update (section 4.9.2.1).

In most cases, this last iteration tends to preserve the features well matched by previous models, and, in general, or it seeks further improvements or it leaves the waveforms unchanged. However, considering the whole dataset, it is possible that the fit between data and synthetic in a given seismogram decreases for model $\mathbf{m}^5$ (e.g., T component in Fig. 4.36). Moreover, sometimes, fitting a part of the seismogram may worsen another part (e.g., T trace in Figs. 4.26 and 4.33, R in Fig. 4.28, or Z in Fig. 4.35), as already noted for previous iterations (section 4.9.1). In this case, the previously selected window shorten or even is rejected, or the corresponding traveltime difference increases (the bar becomes more red or more blue).

It is worth noting that there are several examples in which the initial model $\mathbf{m}^0$ already gives a quite good fit of the data (e.g., Figs. 4.29, 4.37), thus $\mathbf{m}^5$ performs only slight adjustments. This is in agreement with the results of chapter 3, that shows how the 3D model for central Italy used for L'Aquila event simulations is able to reproduce many features missed by 1D models. Moreover, it supports our choice of considering this 3D model as the starting point of the tomographic inversion.

The first part of the seismograms is considered to be mainly influenced by source effects (see chapter 3). It is, thus, consistent not to observe remarkable changes in the first seconds of the timeseries, since we do not invert for source parameters. Moreover, also in the updated model $\mathbf{m}^5$, amplitude fitting is still missing, revealing again the need for amplitude tomography and source inversion.

Analysing waveform comparison for the six considered categories (each component-period pair), one notes that, in general, the misfit tends to be reduced for each of them. However, radial and vertical components, for all the iterations included $\mathbf{m}^5$, seem to have a more similar behaviour compared to transverse component. Moreover, as expected, the windows selected in the period range 6-20 s generally cover a larger part of the seismograms (e.g., Figs. 4.27, 4.29, 4.32, 4.33, 4.35) and they are higher in number with respect to those for 2-20 s. The number of measurement windows, that have been used in the last tomographic iteration $\mathbf{m}^5$, is reported in Table 4.3 both for each category and for the whole dataset.

One should remark, as a conclusion, that model $\mathbf{m}^5$ is still not able to fit all the phases in the observed seismograms. Moreover, for the same source-receiver pair, it happens, as shown, that a component may present relevant improvements, whereas the fit on other components worsen; same thing for different periods. In fact, our inversion involves any type of phases on the three components of ground motion and for two period bands. Thus, a single iteration can not result in a simultaneous fit of all the picks in the seismograms, but successive updates are required to account for all the complexities.

	2-20 s	6-20 s	Total
Radial (R)	1130	1907	3037
Transverse (T)	1055	1707	2762
Vertical (Z)	1280	2044	3324
Total	3465	5658	9123

Table 4.3: Measurement windows for model $\mathbf{m}^5$ - The number of measurement time windows, selected in the last iteration $\mathbf{m}^5$, are reported for each component (R, T or Z) and period range (2-20 s or 6-20 s), i.e., for each category. The total for every component or for every period range or for the whole dataset is also reported.

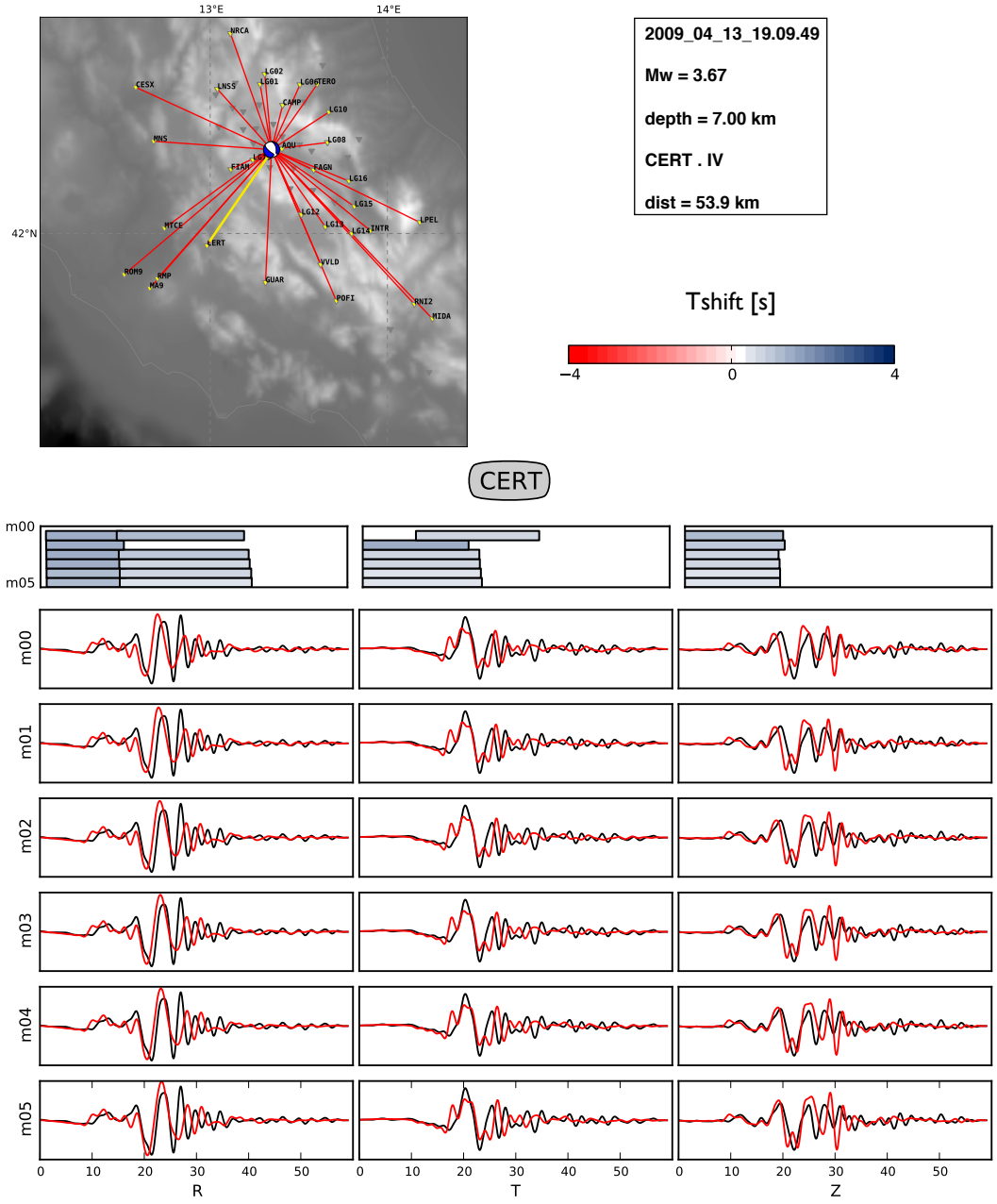


Figure 4.26: Updated waveform comparisons - The map and the box on its right show the considered event and the recording station. At the bottom of the picture, the three components of the recorded data (black) are compared to the synthetics (red) for the initial model $\mathbf{m}^0$ and for each of the updated models $\mathbf{m}^1$-$\mathbf{m}^5$ (see the labels on the left). The timeseries are in displacement (in m) and filtered between 2-20 s. The bars in the boxes (above each seismogram component) indicate, for each model iteration (from $\mathbf{m}^0$ to $\mathbf{m}^5$), the time windows selected in the timeseries, and their colors define the values of the traveltime difference in the window. Based on the Tshift color scale, when color approaches white, the time shift approaches zero and the fit improves.

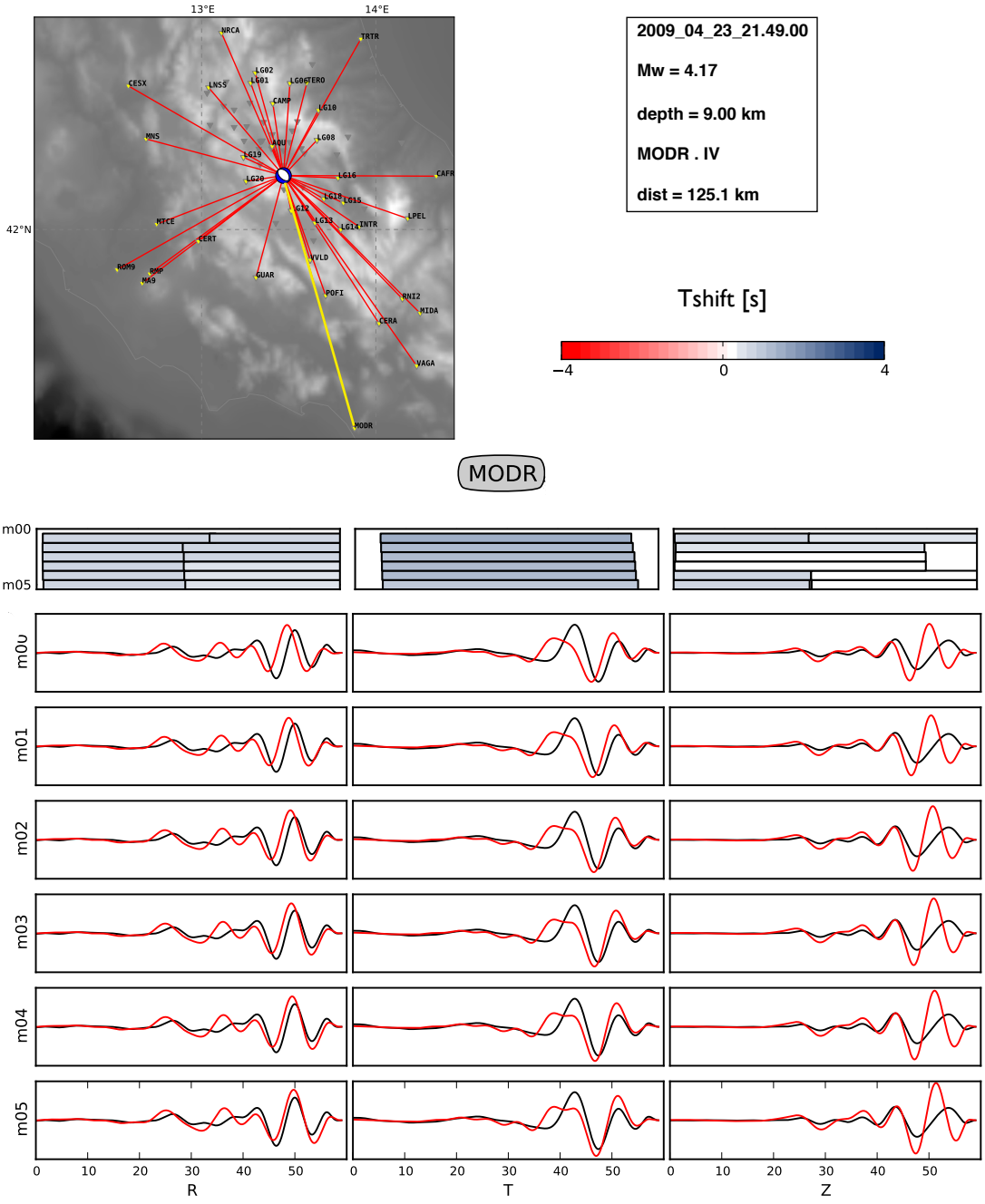


Figure 4.27: Updated waveform comparisons - Same as Figure 4.26, but for different event and station, and for the period range 6-20 s.

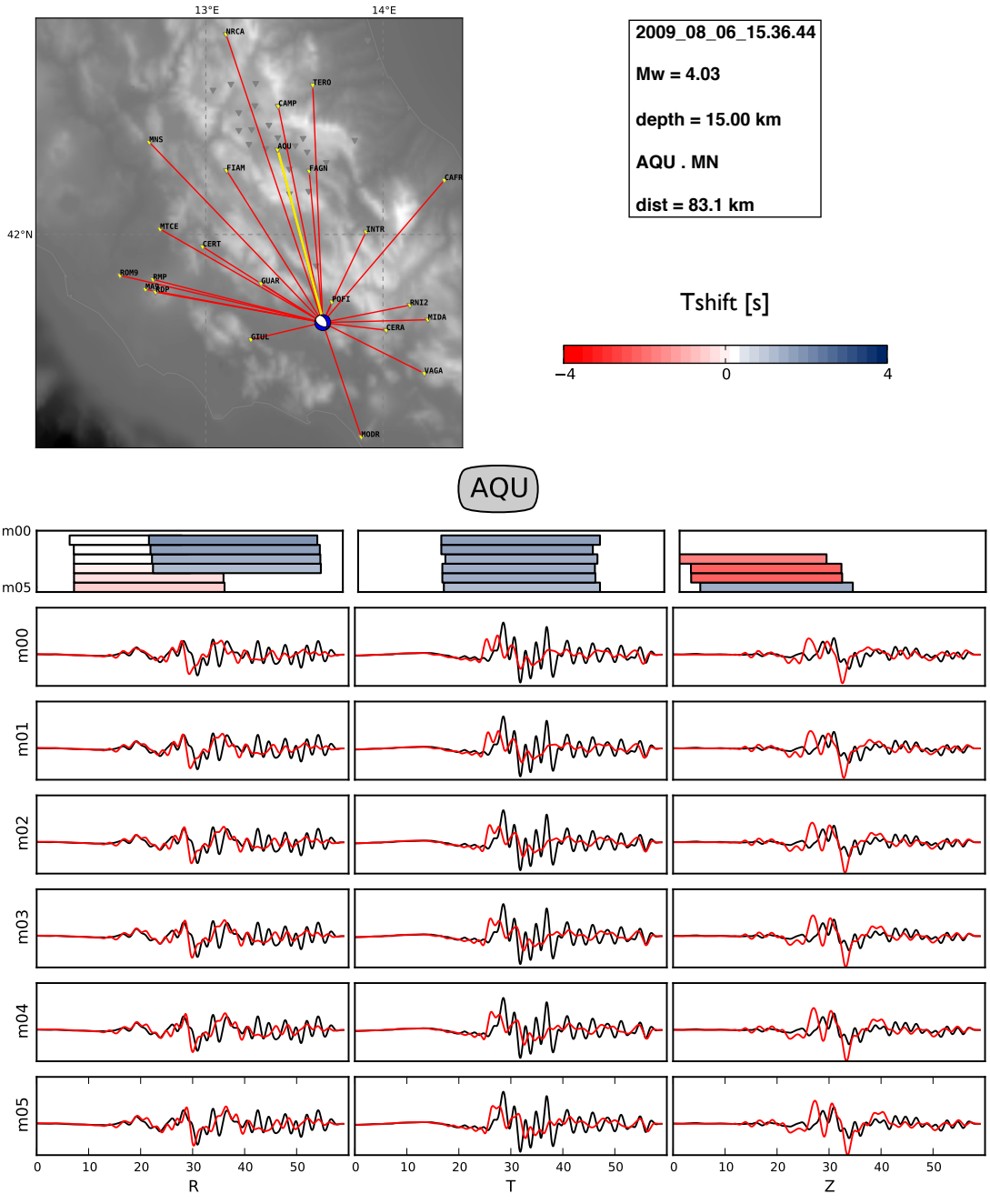


Figure 4.28: Updated waveform comparisons - Same as Figure 4.26, but for different event and station.

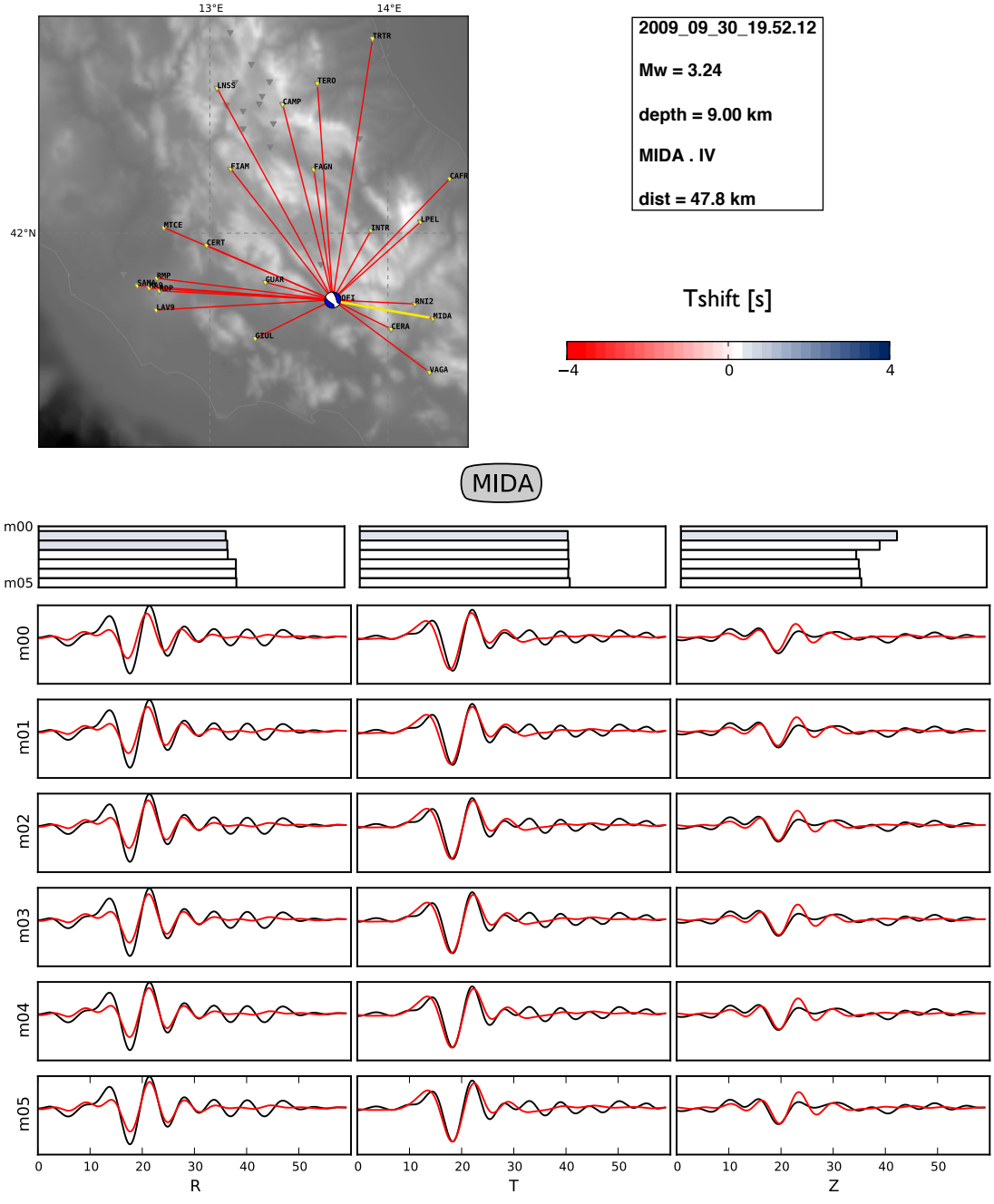


Figure 4.29: Updated waveform comparisons - Same as Figure 4.26, but for different event and station, and for the period range 6-20 s.

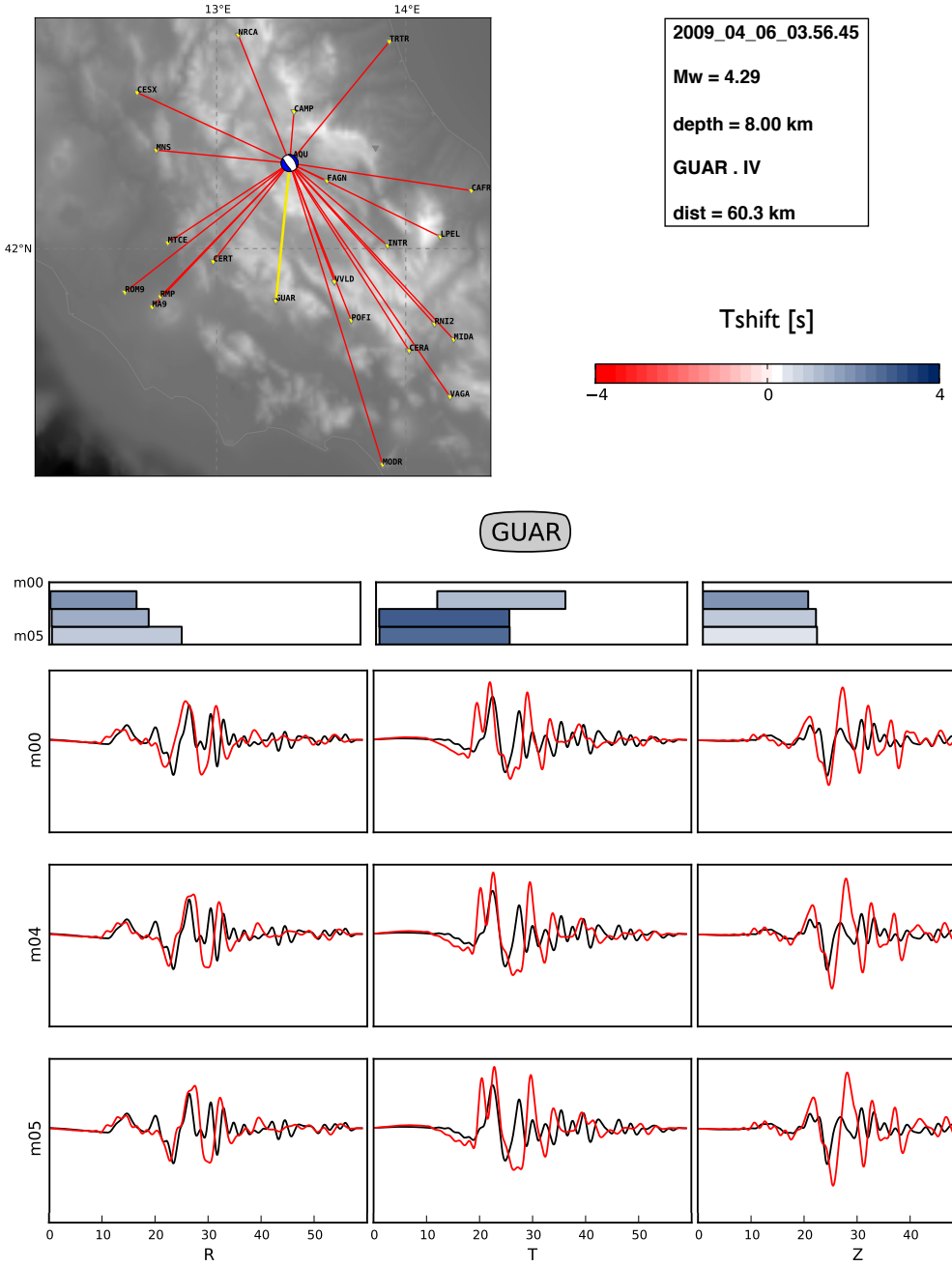


Figure 4.30: Updated waveform comparisons - Same as Figure 4.26, but for different event and station. Moreover, only $\mathbf{m}^0$ and the updated models $\mathbf{m}^4$ and $\mathbf{m}^5$ are shown.

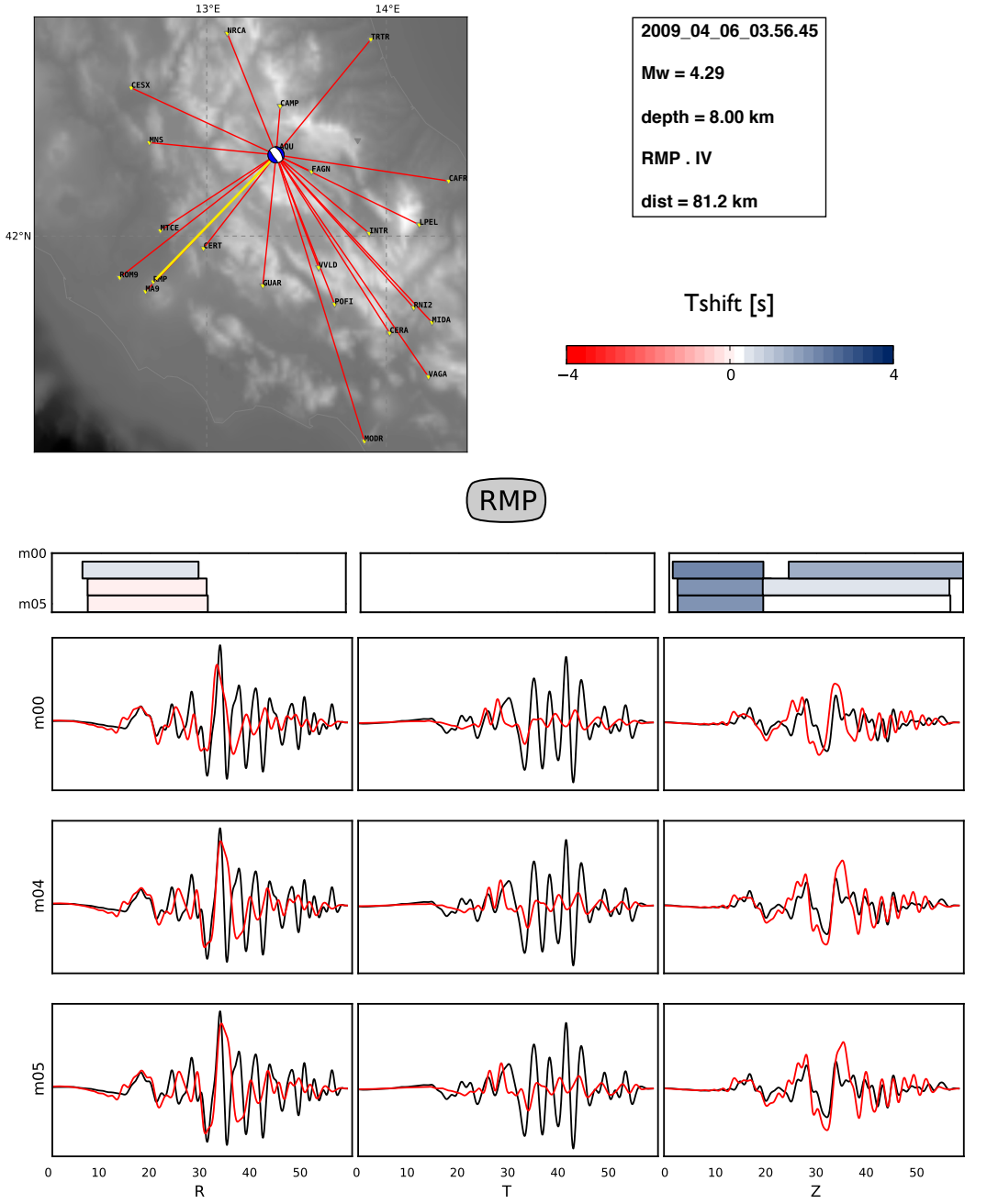


Figure 4.31: Updated waveform comparisons - Same as Figure 4.30, but for a different station.

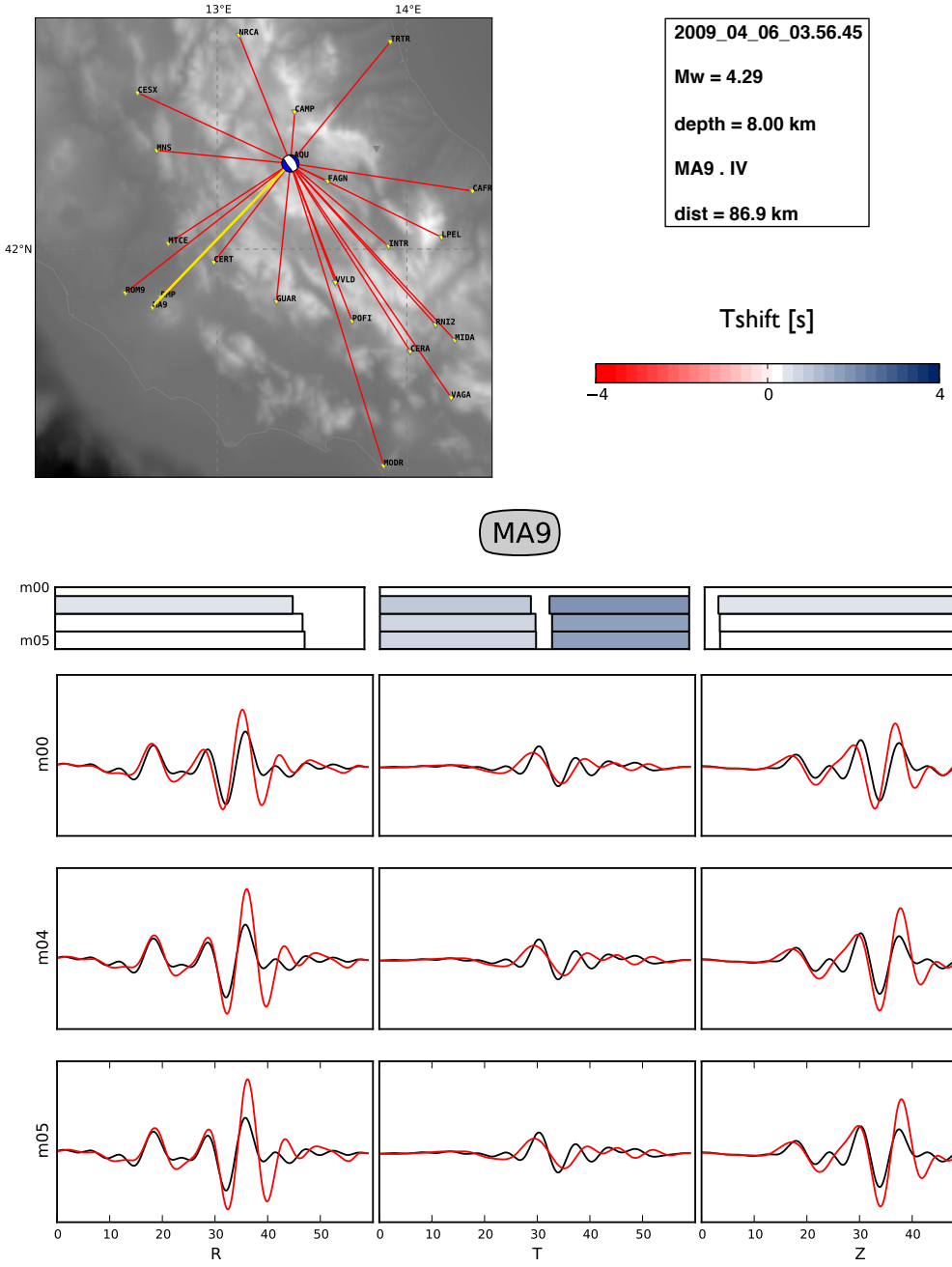


Figure 4.32: Updated waveform comparisons - Same as Figure 4.30, but for a different station, and for the period range 6-20 s.

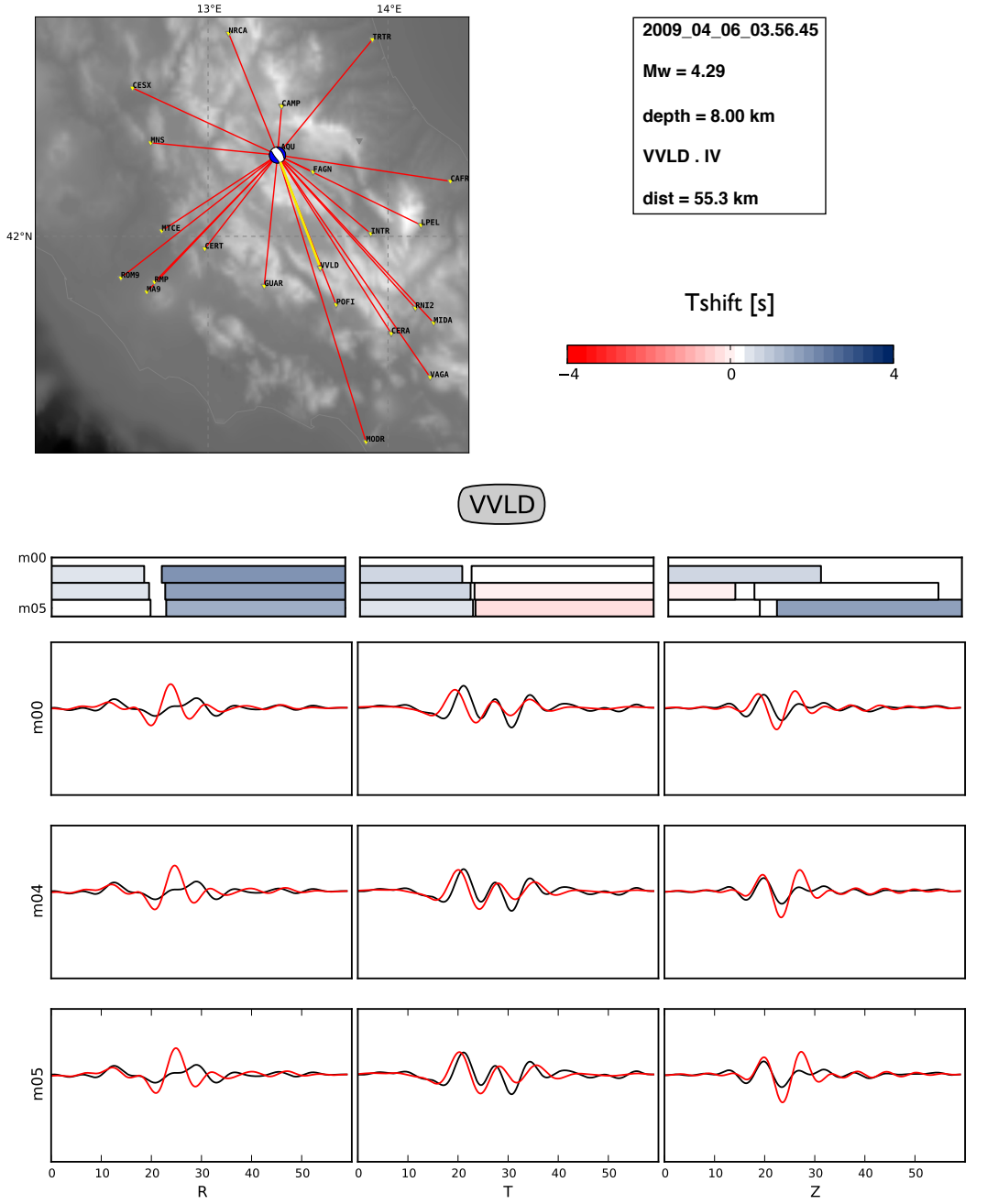


Figure 4.33: Updated waveform comparisons - Same as Figure 4.30, but for a different station, and for the period range 6-20 s.

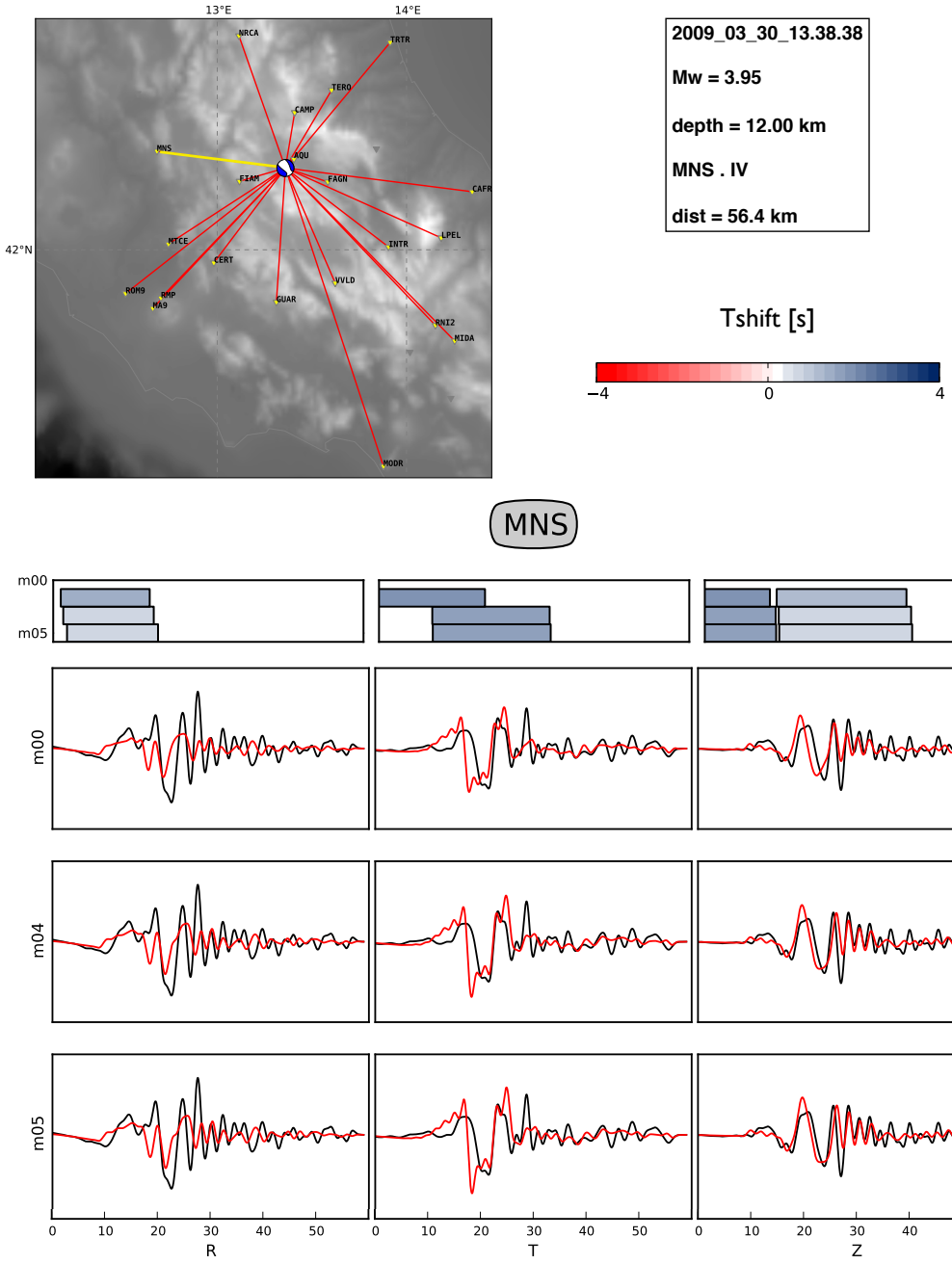


Figure 4.34: Updated waveform comparisons - Same as Figure 4.30, but for different event and station.

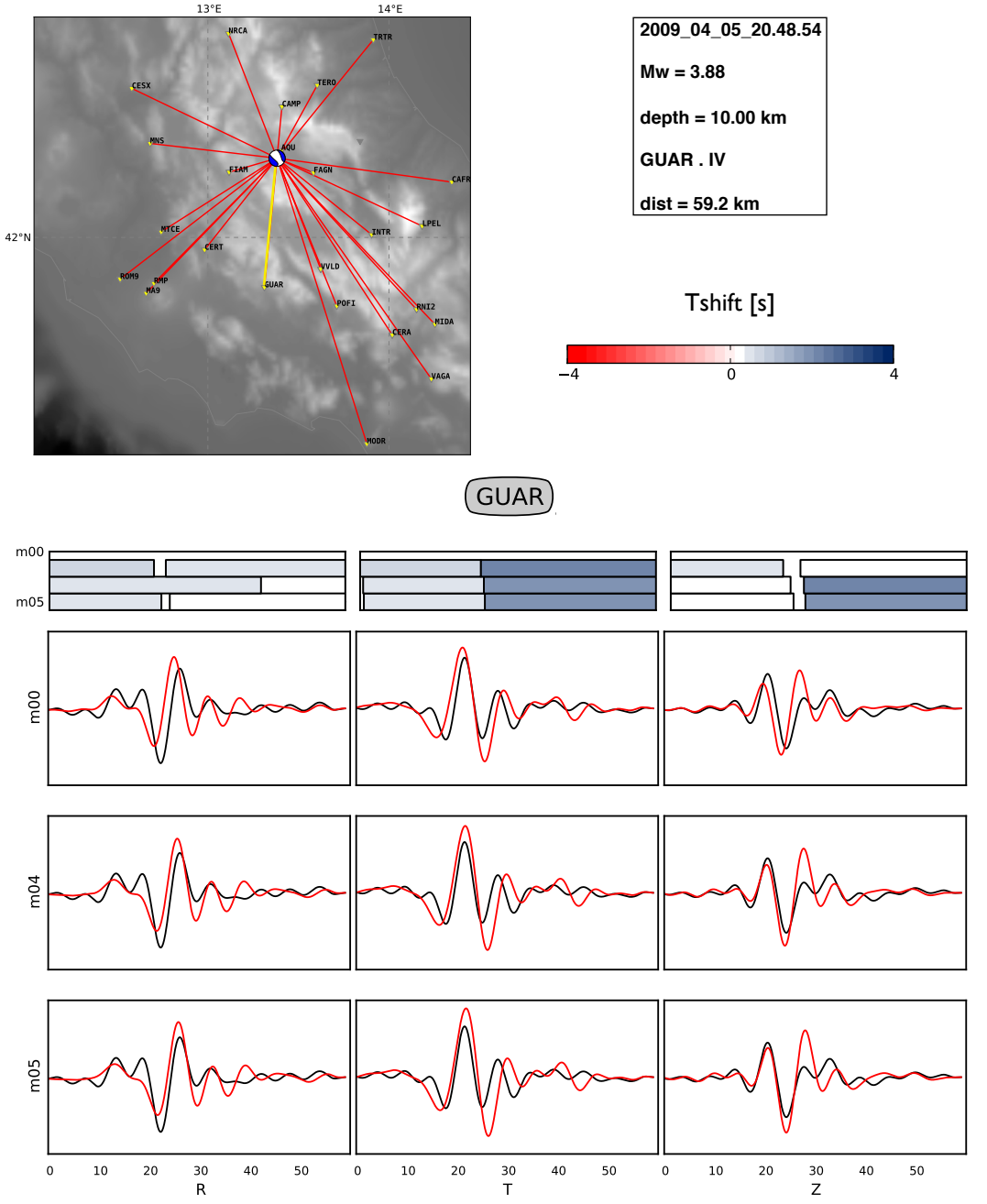


Figure 4.35: Updated waveform comparisons - Same as Figure 4.30, but for a different event, and for the period range 6-20 s.

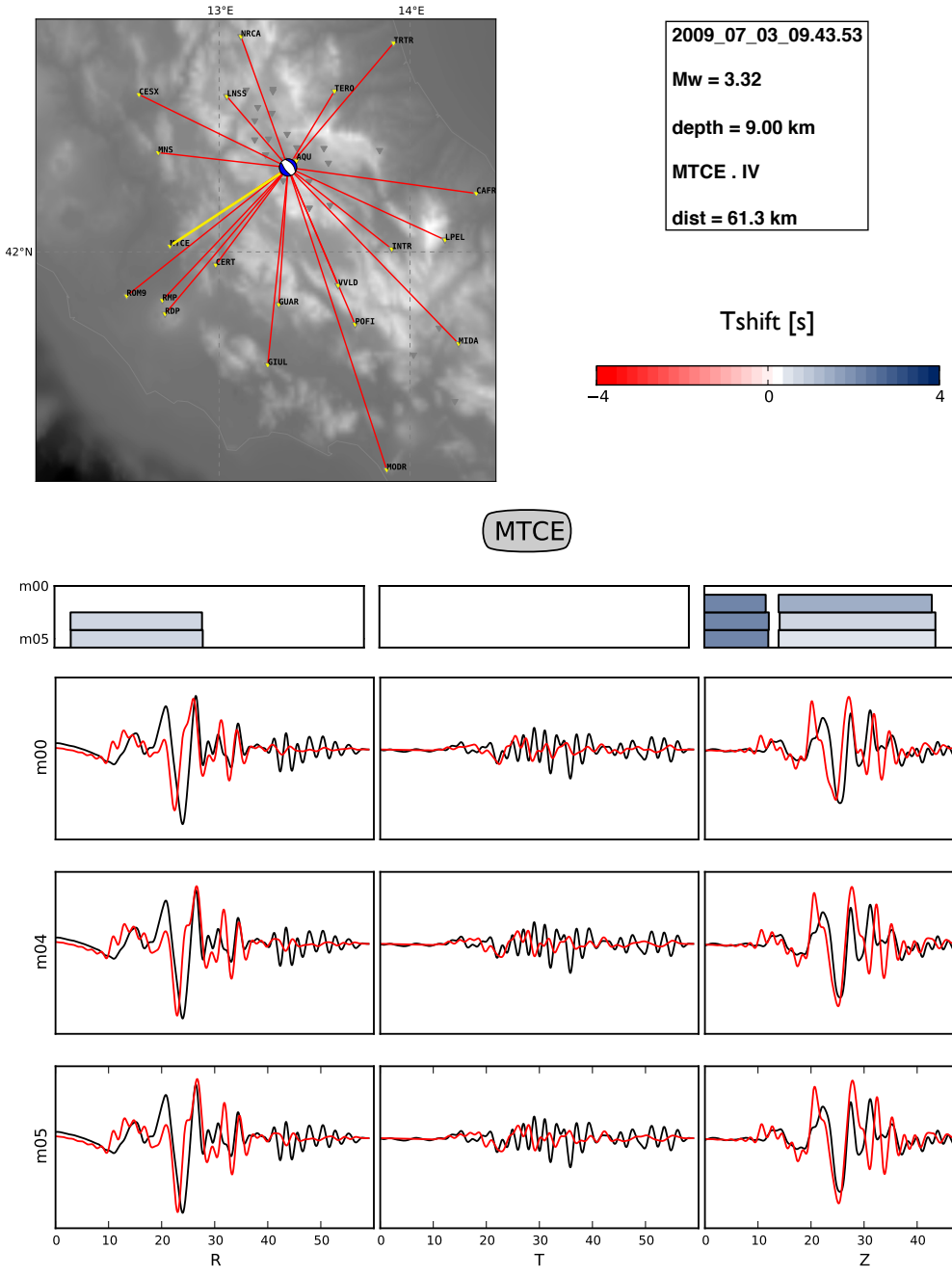


Figure 4.36: Updated waveform comparisons - Same as Figure 4.30, but for different event and station.

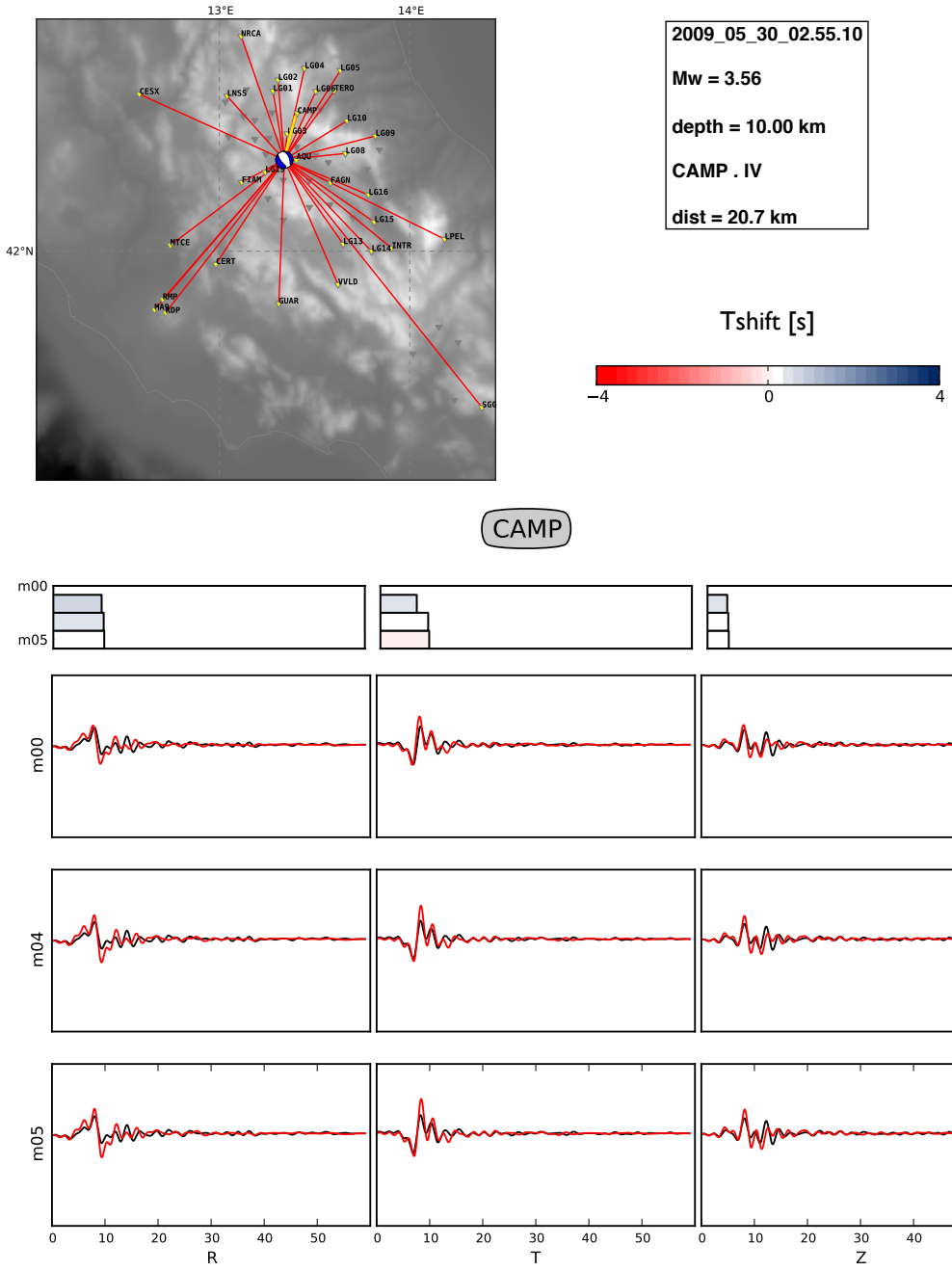


Figure 4.37: Updated waveform comparisons - Same as Figure 4.30, but for different event and station.

4.10 Computational demands

The whole inversion procedure described above is computationally very expensive both in terms of memory and CPU time requirements. To tackle this issue we took advantage of the HPC resources at INGV[1], as we did to perform L'Aquila simulations (section 3.4.4). However, in this case, we exploited both the two available 512-core clusters, running our simulations on 256 cores. One of the two clusters, *ELIOS*, has been already described in section 3.4.4. The other, *SELENE*, has 32 compute nodes, each with 2 oct-core AMD Opteron 6136 processors at 2.4 GHz and with 32 GB RAM (512 total cores, 2 GB RAM/core).

For each event considered in the inversion we need to run one forward simulation to obtain the synthetic wavefields for the misfit estimate. Moreover, an adjoint calculation, i.e. one forward and one adjoint simulations, needs to be performed for each earthquake, in order to calculate the corresponding event kernel. Thus, considering N_s events, the procedure requires in general $3N_s$ simulations at each iteration.

In the tomographic inversion for central Italy, as described above, for model $\mathbf{m}^0$ and $\mathbf{m}^4$ forward and adjoint calculations have been performed for all the considered 63 events, for the last model $\mathbf{m}^5$ only the initial forward simulations for the 63 events have been run, and for all the other iteration models ($\mathbf{m}^1$, $\mathbf{m}^2$, $\mathbf{m}^3$) the simulations involved only the subset of 11 events. In addition, five steplength tests have been performed to determine the α values to calculate each of the five improved models. Every steplength test requires 11 forward simulations for the considered 11 events. The total number of simulations for the whole procedure is, thereby, 705.

Using 256 cores of the INGV clusters, one forward simulation with a duration of 60 s requires about 40 minutes. Instead, a pair of forward and adjoint simulations requires about 3 hours, more than three times the time of a forward simulation (as expected, section 2.4) due to writing and input-output operations. As a consequence our inversion procedure required $1.7\text{x}10^5$ CPU hours.

In addition to the simulation hours, one has to consider the time to process the seismograms and the kernels (sum and regularization), and to run `FLEXWIN` and `measure_adj`, i.e. to select the windows and to estimate the measurements, misfit functions and adjoint sources. Figure 4.38 summarizes the time required to perform one iteration using

[1] http://hpc.rm.ingv.it/

11 events or all the 63 events. These times are estimated without considering the waiting time caused by the queueing system of the clusters. As a result, in principle, about 4 days are required for one iteration using 11 events, but about 2 weeks are needed if 63 events are involved. This is the main reason why for the first iterations we preferred to use only a subset of representative events, although considering all the 63 events could provide valuable information and help to increase the convergence of the procedure (section 4.9.2).

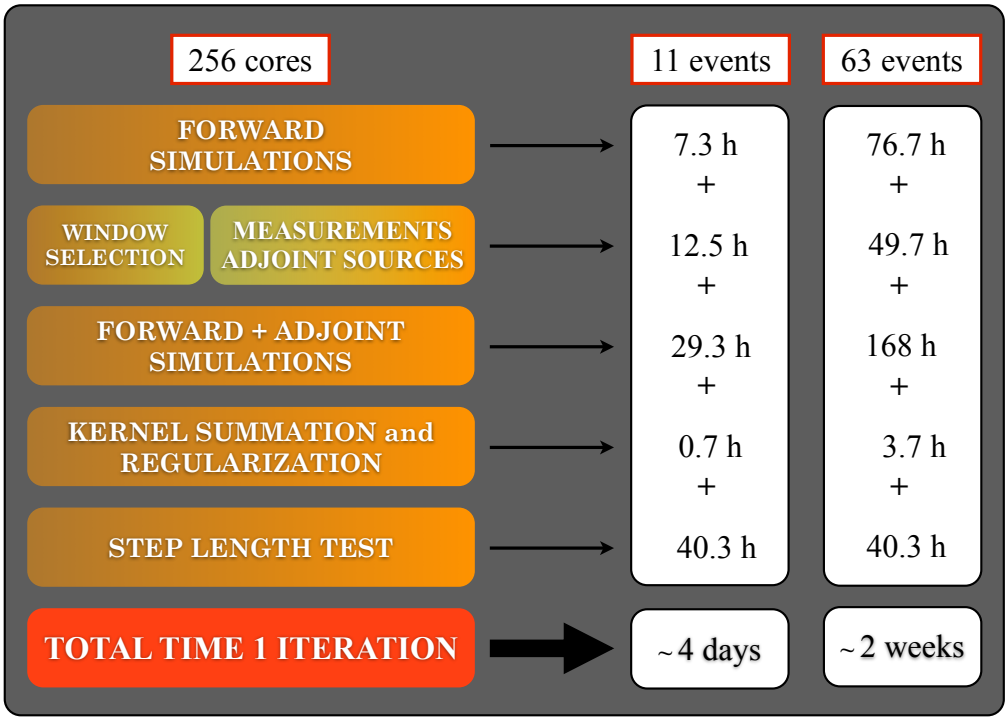


Figure 4.38: Computational requirements - Details on the time required for one iteration of our adjoint tomographic inversion, considering only 11 events or all the 63 events and using 256 cores of the INGV clusters to perform forward and adjoint simulations. The calculated total times involve all the steps of the procedure included the estimate of the total misfit function for the given iteration. Note that the steplength test always involves 11 events (sections 4.7.3 and 4.9), thus the required time is always the same. Moreover, the considered 40.3 hours for this test include both the forward simulations and the measurement estimates for each tested model (see section 4.7.3).

4.11 Discussion and conclusions

In this chapter we presented the results obtained for central Italy after five iterations of a 3D tomographic inversion based upon the SEM and the adjoint method.

The last updated model $\mathbf{m}^5$ does not yet represent the final lithospheric structure of the region. However, successive iterations contribute to reduce the misfit between data and synthetics, and the waveform fitting tends to be progressively improved. Thus, the minimization procedure is properly working in this local context, and we are confident that further iterations, also including source inversions, will allow us to construct a new reference 3D velocity model for central Italy.

Using SEM to simulate synthetic fields for each new iteration, enables us to deal with very accurate waveforms, that account for all the 3D complexities implemented in the numerical technique. Thus, the measured differences between data and synthetics can be only attributed to the discrepancies between the effective real structure and the models. This leads us to perform the inversion based on very accurately constructed kernels, that reflect the sensitivity to the misfit, allowing us to minimize it.

We also benefit from the powerful combination of the SEM with the adjoint method. The paramount advantage is that this technique is based upon misfit kernels, that are sums of the kernels for every single event, and each event kernel, in turn, is obtained by performing only two simulations. These misfit kernels, representing the gradient of the misfit function, highlight where the current model is inadequate and needs to be updated. Thus, firstly, the overall sensitivity to model perturbations, that leads the inversion and is revealed by the misfit kernels, can be obtained by considering only few representative events. This is crucial when a relatively small number of suitable earthquakes is available to describe the considered region, and it is a great advantage with respect to a classical tomography, that, instead, needs to detect thousands of phases. Moreover, the computational burden is dramatically reduced also because the number of simulations required for each iteration depends only on the number of earthquakes. Thus, one can use as many receivers and measurements as possible for the same cost, largely enhancing the accuracy of the inversion. In addition, a tomography based upon 3D models and 3D full-wave simulations, generally too expensive and almost unaffordable using classical techniques, becomes feasible.

We choose to start the tomographic inversion using a 3D tomographic structure for central Italy as initial model. To study the 2009 L'Aquila mainshock (chapter 3) we constructed a 3D model using all the available information on the simulated region. The obtained results revealed that such a 3D representation of the structure is necessary in order to capture many of the features involved in a seismogram, that a 1D model is not able to reproduce. This is the primary reason why we assumed a 3D model, the same used for L'Aquila, as starting point of the inversion. The results of our tomography, then, show that the adopted 3D structural description produces synthetic wavefields that in few cases need only a slight adjustment to fit the data (e.g., Figs. 4.29, 4.37). Thus, in general, using an initial 3D model provides stronger constraints to the inversion, that favour retrieving a very accurate updated model. Moreover, as discussed, using SEM, one can manage the issue of simulating complex 3D wavefields, and the adjoint method makes a 3D inversion feasible. Thus, it is worth trying to solve the inverse problem based on an initial 3D model that incorporates all the valuable information at our disposal.

The main result after five iterations of the steepest descent optimization algorithm is that the misfit between data and synthetics is significantly reduced. For the last iteration $\mathbf{m}^5$, although it is not the definitive model, the total traveltime misfit function is decreased with respect to the initial model $\mathbf{m}^0$ (Fig. 4.25 (a)). Moreover, many arrivals in the seismograms start to be accurately fitted for both considered period ranges, 2-20 s and 6-20 s, and three components of ground motion (Figs. 4.26 - 4.37). Thus, more and more features of the timeseries, captured by the new selected windows, are involved in the inversion, aiming to progressively fit the entire waveforms.

As a consequence, one can obtain successive refinements of structural features in the model, and, possibly, highlight some new characteristics initially missed. The updated model $\mathbf{m}^5$, with respect to $\mathbf{m}^0$, primarily presents reductions of seismic wave velocities that reach a maximum of 4% for v_p and 10% for v_s. Velocity decrease interests most regions of the considered volume (Figs. 4.12 - 4.17), including the main sedimentary basins, that feature high reduction percentages, as well as several mountain reliefs. Only limited areas, instead, have been subjected to velocity increments in the update. These results are an indication of the fact that the initial model was mainly characterized by too high seismic wave speeds, as confirmed by looking at seismogram comparisons (e.g., Figs. 4.26 - 4.37). Synthetics for $\mathbf{m}^0$ mostly feature too early arrivals with respect to

the data; thus, progressively slower models until $\mathbf{m}^5$ tend to shift forward these phases to fit the observed ones. As a consequence, mainly the positive traveltime differences are reduced to reach zero values (e.g., Fig. 4.25 (b)).

Observing the characteristics of model $\mathbf{m}^5$, presented by analysing the six cross-sections in Figures 4.19 - 4.24, one generally notes very slight changes for v_p profile in depth, limiting its variations to a shallow thin layer (about 10 km depth). On the contrary, v_s features a remarkable reduction also in depth (down to about 30 km depth) and its pattern has been sensibly modified compared to $\mathbf{m}^0$. These behaviours of v_p and v_s result in remarkable changes also in the v_p/v_s model for the region. In particular, we observe increments up to 10% of this ratio (that reaches a maximum of 2), mainly evident in the central part of the volume. The obtained high values of v_p/v_s seem to be consistent with recent studies (e.g., Di Luccio et al. (56), 2010; Lucente et al. (120), 2010), which attribute high values of v_p/v_s, mainly related to low values of v_s, to the presence of fluids. However, the increase of the seismic velocity ratio in $\mathbf{m}^5$ may be not perfectly realistic, but affected by the too strong reduction in depth of v_s with respect to v_p. This could be though as a limit of our tomographic inversion, that is based upon compressional-wavespeed kernels K_α with a weak resolution in depth compared to shear-wavespeed kernels K_β. Moreover, the high v_p/v_s values inferred by, e.g., Di Luccio et al. (56) (2010) and Lucente et al. (120) (2010) mainly occur as a consequence of the 2009 L'Aquila mainshock, that has been hypothesized to cause fluid migration. Our tomographic inversion involves earthquakes both before and after L'Aquila event. Thus, we can not properly separate effects on the velocity model related to pre- or post-mainshock seismic properties, although the larger number of aftershocks, compared to foreshocks, could have a higher influence.

In general, the robustness and reliability of most of the observed features in $\mathbf{m}^5$ seem to be confirmed by successive iterations. However, we should be aware in interpreting these structures from a geologic and tectonic stand point. We prefer to stress on how the minimization proceeds and how the model is changing with respect to the initial one, postponing interpretations when a tested definitive model will be obtained. In fact, $\mathbf{m}^5$ itself is still an iteration, and more steps of the inversion procedure are required to obtain a final tomography.

This is evident, for example, observing that some features present in an iteration disappear in the next one and viceversa, indicating that we need a higher resolution

to model them or to consider them spurious. Moreover, observing the seismograms (e.g., Figs. 4.26 - 4.37), although fit improvements are evident, many phases are still not matched for a particular component or period band, and sometimes they can also worsen, while improving the others. This is something expected in an inversion that involves many types of seismic phases for different ground motion components and period ranges. A simultaneous improvement of all phases at all periods is not possible with a single iteration, and during the procedure some fittings may decrease. Only after several steps, one can observe remarkable adjustments in almost all the phases in the seismograms. The southern California adjoint inversion of Tape et al. (189) (2010) required 16 iterations to obtain a final tomographic model, and relevant changes with respect to the initial structure become evident only after numerous iterations.

Comparisons between data and synthetics also highlight the limits of our tomographic inversion. The initial discrepancies between observed and synthetic seismograms have been mainly attributed to errors in the structural parameters. Thus, we decided to firstly perform a traveltime tomography, that primarily reduces the traveltime differences between phases. During five iterations, while phase shifts continue to be minimized, amplitudes and waveform shapes are only slightly adjusted. This reveals that, as expected, the choice of the misfit function is critical in defining which features of the model will be improved in the inversion. A possible way to enhance model updates for central Italy is to perform in the future also an amplitude tomography study. However, we believe that at this point of our inversion a significant improvement will result by performing a source inversion. For each of the considered events we used, at all iterations, the source solutions from the TDMT technique (section 4.3), that is based upon a 1D velocity model. In order to seek full waveform fitting, a refinement of moment tensors and other source parameters is absolutely required. In fact, assuming perfect source solutions, errors on these parameters are mapped into the structure and only a tomographic inversion can not resolve all these inaccuracies. Moreover, using source parameters from an inversion based upon a 3D model would be more consistent with the considered 3D structural models. For these reasons we started performing some preliminary source inversion tests using a semi-automated technique (Liu et al. (119), 2004). It allows moment tensor and earthquake location estimates, based upon SEM and 3D velocity models. The basic steps of the procedure are outlined in Appendix A. This thesis does not involve updated source solutions, mainly due to the computational

costs of the technique (Appendix A). However, the next step for a future work will be to use the improved model $\mathbf{m}^5$ in order to invert source parameters of all the considered events, and then continuing the tomographic iterations.

Other critical choices in the inversion are the preconditioning and smoothing functions. Preconditioners represent further constraints in the minimization procedure. We choose a function that has an expression similar to the misfit kernels (section 4.7.2) and it properly results in an increment of the convergence (Fig. 4.7 (d)). However, other choices were possible (e.g., Fichtner et al. (75), 2009b; Tape et al. (189), 2010) and it would be interesting to study the effects of different preconditioners. The smoothing, as well, needs to be carefully applied, since its scalelength defines the resolution of the problem. In order to establish a suitable value in our case, we performed some preliminary tests. We prefer to be conservative and to use a smoothing with a scalelength quite larger than the minimum wavelength resolved, at least for these first iterations.

In a tomographic inversion it is desirable to use a set of earthquakes that properly samples the considered medium, since the resolution scalelength primarily depends on data coverage. In our case, we tried to select events as much homogeneously distributed as possible, both horizontally and in depth. However, most of them are clustered around the L'Aquila mainshock, as expected. Thus, the sensitivity of the resulting kernels is mainly concentrated in the central part of the volume, and the strongest variations of the updated model $\mathbf{m}^5$ occur in that region. In the future, more events with different location could be included in the inversion, in order to constrain information for poorly sampled areas. Anyway, events located near to each other can be useful, since at short periods they can have very different waveforms proving different information. Moreover, at longer periods, one could use discrepancies between event kernels for similar sources to highlight bad sources (e.g., Tape (186), 2009).

In order to define when the iterative procedure should be concluded, a first indication is that the misfit function starts oscillating between iteration and/or its value approaches 1, that means the data have been fitted within one standard deviation (section 2.3.4.2). Moreover, one should suspend iterating when becomes evident that other factors influence waveform discrepancy, no longer reduced by the tomographic inversion. Our results seem to suggest a similar condition. Although further slight adjustments may be obtained continuing iterating using, for example, all the 63 events, at

this point, an improvement of source parameters seems necessary. This is the primary reason why, for now, we decided to stop the procedure at iteration $\mathbf{m}^5$.

Looking forward, when a final model is obtained, one can firstly proceed by accurately assessing the misfit behaviour. To this end, together with traveltime misfit, a full-waveform misfit function could be considered (e.g., Tape et al. (189), 2010), that represents the straight difference between data and synthetic seismograms. Moreover, a resolution analysis will be needed, and a possibility is to calculate the volumetric sensitivity of all measurements (e.g., Chen et al. (36), 2007; Tape et al. (189), 2010). In addition, one could analyse individual sensitivity kernels for specific interesting phases, and other events, not involved in the inversion, could be used to test the new model (e.g., Tape et al. (189), 2010). A detailed interpretation of the definitive model will be finally allowed.

To conclude, the last iteration $\mathbf{m}^5$, that we obtained for the tomographic inversion in central Italy, provides us valuable results, that show interesting variations with respect to the initial model. It represents an improved description of the structure, characterized by a reduced misfit and an updated data fitting. We consider these results as a good starting point for future improvements, that will lead to a new final 3D model for central Italy lithosphere.

Chapter 5

Conclusions

In this work we started by exploiting the powerful capabilities of a spectral-element numerical method (SEM) in order to simulate, at relatively high frequencies (up to 0.5 Hz), the wavefield generated in central Italy by the 2009 M_w 6.3 L'Aquila earthquake (chapter 3). The results of this study represent pillars for the final aim of our work, which consists in performing 3D full waveform tomography of the central Italy lithosphere, based on the combination of the SEM and the adjoint method (chapter 4).

The results presented in chapter 3 show that, by using a finite fault representation and a 3D structural model with topography, we are able to realistically replicate the waveform data associated to the event. This is one of the first studies that takes into account both the complexities related to a source of finite size and those arising from a realistic 3D structure, and that compares the corresponding synthetics with the data. The key results and aspects, discussed in detail in section 3.9, are summarized in what follows.

- A 3D representation of central Italy structure, that incorporates all the possible (known) heterogeneities in the region, included topography, is essential. This allows us to fit all three ground motion components of the observed data, between 0.02 and 0.5 Hz, with a good level of accuracy.

- Considering a finite fault model and only 1D velocity profiles, with and without topography, focuses the attention on the necessity of including a detailed topographic description to model many of the features in the seismogram coda. However, just the topography is not sufficient to reproduce all the complexities in

the wavefields, supporting again the need for a 3D model. In addition, using the same 1D profile adopted by Cirella et al. (42) (2009) to infer the source model now considered, we obtain synthetics in agreement with their results. This means that we properly reproduce their kinematic fault description, although slight discrepancies arise probably because we adopted a different source time function.

- Using the constructed 3D model, we are able to capture the effect, on the waveforms, of features not included in Cirella et al. (42) (2009)'s source inversion, such as the topography and low wavespeed basins in the region. Moreover, also the data recorded at stations not accounted for by Cirella et al. (42) (2009) show discrete matching with synthetics. These aspects suggest that the considered structure description already contains valuable information to model seismic wave propagation, although improvements are required.

- The effects of the topographic ridges and of the low wave velocity sedimentary basins in central Italy mainly affect the coda of the observed seismograms. They generally cause amplifications, multiple scattering and prolongation of the seismic shaking, also because the topography enhances and complicates the resonance in the basins. Testing different structure descriptions (1D and 3D), we infer that only a 3D model (with topography) is able to highlight all these features. On the other hand, the source effect is prominent in the first part of the seismograms and the three considered models (both 1D and 3D) have a similar behaviour in these initial portions. This is consistent with the fact that the simulations for all these cases involve substantially the same finite fault model.

- Despite our synthetics for finite fault and 3D structure models are able to capture some important characteristics of the observed seismograms, they do not still account for many arrivals and often overestimate the data. This firstly suggests that the considered structure model needs to be refined, since, likely, several features of the real structure are poorly constrained or even neglected. In particular, the sedimentary basins in central Italy, implemented using the v_s30 model, would require a more detailed description, explicitly incorporating their shape and depth. Moreover, the amplifications in the synthetics, with respect to data, could be ascribed to an overestimate of the on-fault slip. This kinematic parameter, being

derived from a source inversion based on a 1D profile, seems poorly consistent with the 3D velocity model used in the simulations. This highlights an intrinsic incompatibility between the considered source and structure models, unsolvable in this work. Possible finite source inversions based on 3D structure models have been very recently presented (Trasatti et al. (196), 2011; Volpe et al. (203), 2012) and could be useful for our future analyses.

- The accuracy of our forward modeling for waves propagating in central Italy is presently limited up to a frequency of ∼0.5 Hz (i.e., down to a period of ∼2 s). This is primarily due to the limited accuracy of the available source and structure models, that do not still reproduce in detail features evident in the seismograms at higher frequencies. This offers the possibility of future improvements.

- The 3D structure model, together with the finite fault representation, significantly improve the agreement between observed and synthetic estimates of peak ground velocity (PGV) for L'Aquila event. With respect to the 1D models, the effect of the low wavespeed basins of increasing PGV is highlighted. Moreover, evidences of the topographic ridges and of source directivity, already shown by the 1D profile with topography, are enhanced by the 3D model. This supports once again the need for a 3D description of the structure, also fundamental for seismic hazard assessment, on which our synthetic PGV maps already suggest useful indications.

- Finally, we used the SEM to simulate the complex ground motion generated by L'Aquila event. This guaranties a strong numerical accuracy, that will benefit and encourage also higher resolution simulations. Moreover, the results confirm the power of the method in implementing all the complexities of a kinematic source together with a realistic regional structure.

The results on L'Aquila event proved, thereby, that the available 3D structure model for central Italy already reproduces many features that influence seismic wave propagation in the region. Thus, this model could be used to invert for source parameters of other events in the area, in order to construct 3D source solutions presently not available (as suggested by Komatitsch et al. (109), 2004). However, we also noticed numerous remaining discrepancies comparing the data and the 3D synthetics. Hence, the primary interest of the second phase of our work has been to improve the initial 3D

velocity model for the region, postponing the source inversion to when a more accurate structure become available (Appendix A). Thus, using SEM combined with the adjoint method and the steepest descent gradient-based minimization algorithm, we performed five iterations of the inversion procedure to update the central Italy tomographic model. In what follows we summarize the key points and results, largely discussed in section 4.11.

- The initial 3D velocity model, labeled $\mathbf{m}^0$, is already able to fit some observed arrivals in three-component seismograms for both considered period ranges, 2-20 s and 6-20 s. Thus, the results inferred for L'Aquila event are enhanced by other earthquakes and stations in the region. This supports our choice of starting the inversion from a 3D model, instead of a simpler 1D profile, corroborated by the feasibility of a 3D adjoint tomography.

- Successive iterations of the inversion gradually reduce the total traveltime misfit function between data and synthetics. Thus, progressively updated models are obtained and the corresponding synthetics tend to better fit many arrivals in the observed seismograms. This suggests that the inversion procedure is properly working in view of determining a final model that minimizes the misfit function.

- Only selected windows of the timeseries are involved in the inversion, namely the ones within which data and synthetics have a discrete agreement and suitable amplitudes. However, proceeding with iterations, the matching improves and more phases are fitted and then included, increasing the width and/or number of the selected windows. This will aim progressively to full waveform fitting.

- After five iterations of the inversion we obtain a further updated model for the region, labeled $\mathbf{m}^5$. Despite this is not the new final tomographic model, it provides a structure representation of central Italy once again improved with respect to $\mathbf{m}^0$. In particular, using $\mathbf{m}^5$, the total traveltime misfit function between data and synthetics is further reduced. Thus, the agreement between data and synthetics continues increasing generally for all the components and considered period ranges.

- All the performed iterations mainly tend to reduce the velocities in the initial 3D model $\mathbf{m}^0$. In particular, several regions of model $\mathbf{m}^5$ feature a velocity decrease,

including sedimentary basins and also many topographic reliefs in central Italy. Velocity reductions from $\mathbf{m}^0$ reach maximum values of 4% for v_p and 10% for v_s, and they are confined in a shallow layer for v_p ($\sim$10 km depth), whereas they extend deeper for v_s (also down to $\sim$30 km depth). Only restricted areas are, instead, interested by an increment of the velocities and the percentages are lower ($\sim$1.2% maximum for v_p and v_s in $\mathbf{m}^5$). This pattern of the velocity variations, that is continuous during successive iterations, also results in an increase of v_p/v_s with respect to $\mathbf{m}^0$, mainly in depth in the central part of the volume. For $\mathbf{m}^5$ the maximum increment is about 10% (for a maximum v_p/v_s=2). This effect could be also related to poorly resolved variations in depth of v_p, with respect to v_s, and further iterations are required to assess the reliability of these features.

- In general, while fitting more arrivals, the resolution of the updated models increases and new features can be imaged. However, presently, they should be interpreted with caution, since $\mathbf{m}^5$ is not the definitive model, and more iterations are needed, that could significantly change the characteristics of the structure (as shown by Tape et al. (189) (2010)). We prefer to focus on how the inversion is proceeding, leaving geologic and tectonic interpretations to when the final model is retrieved. Nevertheless, a first insight on what structural features produce specific phase arrivals may be obtained, also at this stage, by constructing the sensitivity kernels for these pulses (e.g., Tape et al. (189), 2010).

- Despite the agreement tend to increase, many of the complexities in the observed seismograms are still not fitted by $\mathbf{m}^5$, particularly when the higher frequencies are included (range 2-20 s). This confirms once again that also the last model presently obtained requires further improvements. In general, an adjoint 3D tomography needs to resort to a gradient-based algorithm and multiple iterations are necessary. Moreover, in our case the inversion involves different periods and waveform components, thus one should expect that a significant improvement in all the categories can occur only after several steps of the procedure.

- We consider a traveltime misfit function, thus mainly the shift between observed and synthetic phase arrivals is minimized. Only slight adjustments are instead evident for the amplitudes and the general shape of the waveforms. Performing also an amplitude tomography may contribute to reduce these discrepancies.

However, starting from model $\mathbf{m}^5$, we believe that a more significant contribution to improve waveform modeling at that point would be provided by a source inversion. Thus, the next step in future developments of this work will be to invert for the source parameters of the considered events, based upon the last 3D updated model (see Appendix A for a preliminary outline).

- The events used in the inversion are mainly foreshocks and aftershocks of the 2009 L'Aquila earthquake. Thus, they are mostly clustered around the mainshock and the sensibility of the inversion is higher in the central part of the considered volume. Although these events may provide different information, sharpening thereby the overall sensitivity, considering, in the future, other events distributed in the volume can help to constrain poorly sampled areas. Moreover, since the resolution of the model increases through successive iterations, also smaller events (M_l <3) could be considered, that are large in number and involve information at higher frequencies. However, this highlights again the need for a source inversion, to determine the solutions not available for these events.

- Our tomographic inversion can exploit the combined power of SEM and adjoint method. The numerical technique assures very accurate full-wave simulations with reduced computational costs. In addition, the adjoint method allows us to perform a 3D tomographic inversion that requires, at each iteration, only two simulations for every event kernel. Thus, one can consider as many measurements as possible and also a relatively small subset of well-suited events can provide the essential data for the inversion.

- Many other assessment analyses would be possible, such as using other events in the region to test the retrieved model, a misfit analysis based on different misfit functions or also a resolution analysis (e.g., Tape et al. (189), 2010). However, we consider more significant performing these studies when a final model is obtained.

- Finally, the promising results obtained with $\mathbf{m}^0$ and the improvement provided by $\mathbf{m}^5$, suggest that it would be interesting to review the study of L'Aquila mainshock using $\mathbf{m}^5$, or better the final improved model (when available). Using a very accurate structure description may allow us to reproduce in more detail the complicated wavefield of such moderate event. This can be paramount for

seismic hazard analyses, also of other larger events. Moreover, it could be useful to produce a new source model for L'Aquila mainshock using, e.g., Cirella et al. (42) (2009)'s inversion technique and the improved 3D structure (instead of a 1D profile). The adjoint methods (combined with SEM), as well, provide important results to address this issue (e.g., Kim et al. (97), 2011). This would contribute to solve the incompatibility between source and structure models highlighted in the L'Aquila event study. In addition, it offers the opportunity of a 3D kinematic source inversion, also for larger earthquakes.

As a conclusion, our studies, both on L'Aquila event and on the tomography of central Italy, provide valuable results concerning forward and inverse modeling of seismic waves in the region. Moreover, they lend themselves very well for further improvements, aiming to reproduce in detail the wavefield generated by large events and to finally retrieve a new reference 3D structure model for central Italy or elsewhere. Interesting future developments can arise delving our analyses with further data and models.

Appendix A

Source inversion

In our tomographic inversion for central Italy we use events approximated as point sources, with moment tensor solutions obtained by a Time Domain Moment Tensor (TDMT) technique, that relies on a 1D velocity profile (section 4.3). The results show that a refinement of the source parameters could be important at this point, before continuing iterating for the structure model (as, e.g., in Tape et al. (189), 2010). An inversion of the source parameters based upon a 3D velocity model can be performed using, e.g., a procedure presented by Liu et al. (119) (2004) and implemented in their code CMT3D. This technique allows one to invert for centroid moment tensor and event location by numerically calculating the Fréchet derivatives with respect to the considered source parameters and then minimizing a waveform misfit function between data and synthetics (section A.1). The wavefield simulations are performed using SEM.

The tomographic inversion alone has large computational requirements (section 4.10). Liu et al. (119) (2004)'s procedure would require several additional simulations for each event (at least 6 - 9, depending on how many source parameters are considered; section A.2). Thus, in this thesis we do not perform source inversions, focusing our attention only on the structure improvement, with fixed source parameters at all iterations. However, our next objective after this work will be to perform CMT inversions for all the considered events, starting from the available source solutions and using the last updated 3D model $\mathbf{m}^5$. Hence, in what follows I will briefly summarize the theory of moment tensor inversion used by Liu et al. (119) (2004) (section A.1). Moreover, in section A.2 I will outline the steps required to perform the inversion using CMT3D, and a very preliminary example is shown using one of the considered events.

A.1 Basic theory

Following Q. Liu[1], one starts by defining a waveform misfit function between the data and the synthetics for a given model $\mathbf{m}$ (cfr. eq. 2.47)

$$F(\mathbf{m}) = \frac{1}{2} \sum_{i=rp} \int W_{rp}(t) \left[s_i(\mathbf{x}_r, t, \mathbf{m}) - d_i(\mathbf{x}_r, t) \right]^2 \mathrm{d}t \,, \tag{A.1}$$

where $\{s_i(t, \mathbf{m}); i = 1, ..., N\}$ and $\{d_i(t); i = 1, ..., N\}$ denote processed synthetics and data, respectively, $\mathbf{m} = \{m_j; j = 1, \ldots, M\}$ is the source model vector, and W_{rp} combines the taper function with weighting factor for the p-th window at r-th receiver.

The first and second derivatives of the misfit function A.1, i.e. its gradient and Hessian (section 2.3.1), at a given initial model $\mathbf{m}^0$ are expressed by

$$\frac{\partial F(\mathbf{m}^0)}{\partial m_j} = \sum_{i=rp} \int W_{rp}(t) \left[s_i(\mathbf{x}_r, t, \mathbf{m}^0) - d_i(\mathbf{x}_r, t) \right] \frac{\partial s_i(\mathbf{x}_r, t, \mathbf{m}^0)}{\partial m_j} \, \mathrm{d}t \,, \tag{A.2}$$

$$\frac{\partial^2 F(\mathbf{m}^0)}{\partial m_j \, \partial m_k} \approx \sum_{i=rp} \int W_{rp}(t) \, \frac{\partial s_i(\mathbf{x}_r, t, \mathbf{m}^0)}{\partial m_j} \, \frac{\partial s_i(\mathbf{x}_r, t, \mathbf{m}^0)}{\partial m_k} \, \mathrm{d}t \,, \tag{A.3}$$

where in A.3 we considered the approximate Hessian by neglecting the second order derivative terms (section 2.3.4).

As in a general inverse problem, one needs to find a model $\mathbf{m}$ that minimizes the misfit function A.1. Thus, the gradient of F evaluated at $\mathbf{m}$ should be equal to zero, that leads us to the quasi-Newton problem that has to be solved [2]:

$$\frac{\partial F(\mathbf{m}^0)}{\partial m_j} \; + \; \frac{\partial^2 F(\mathbf{m}^0)}{\partial m_j \, \partial m_k} \, (m_k - m_k^0) \; = \; 0 \,. \tag{A.4}$$

Hence, given equations A.2 - A.3, in order to solve the inverse problem one needs to compute the Fréchet derivatives of the synthetics with respect to the source model parameters, $\frac{\partial s_i(\mathbf{m}^0)}{\partial m_j}$.

In general, the synthetics can be expressed by a relationship linearized with respect to an initial source model $\mathbf{m}^0$

$$s_i(t, \mathbf{m}) = s_i(t, \mathbf{m}^0) \; + \; \sum_{j=1}^{M} \frac{\partial s_i(t, \mathbf{m}^0)}{\partial m_j} \, (m_j - m_j^0) \,. \tag{A.5}$$

[1] http://www.geodynamics.org/svn/cig/seismo/3D/GRD_CMT3D/cmt3d/readme/readme.pdf

[2] Everything becomes as in section 2.3.1 by substituting $\mathbf{m}^0 \rightarrow \mathbf{m}$, $\mathbf{m} \rightarrow \mathbf{m} + \delta\mathbf{m}$ and $\mathbf{m} - \mathbf{m}^0 \rightarrow \delta\mathbf{m}$.

If the source model vector $\mathbf{m}$ contains only the six moment tensor components M_j, the synthetics can be represented as linear combinations of the Fréchet derivatives with respect to these parameters:

$$s_i(t, \mathbf{m}) = \sum_{j=1}^{6} \frac{\partial s_i(t, \mathbf{m}^0)}{\partial M_j} M_j \,. \tag{A.6}$$

In this case the inversion is linear (provided a double-couple constraint is also neglected) and the Fréchet derivatives required to solve A.4 are given by

$$\frac{\partial s_i(t, \mathbf{m}^0)}{\partial M_j} = \frac{s_i(t, M_j^0)}{M_j^0} \,, \tag{A.7}$$

where $s_i(t, M_j^0)$ is the synthetic wavefield simulated by setting all the initial moment tensor components equal to zero except the j-th. On the contrary, for the other point source parameters, such as depth, latitude and longitude, the inversion becomes non linear, equation A.5 holds and the Fréchet derivatives are calculated using a finite-difference formula

$$\frac{\partial s_i(t, \mathbf{m}^0)}{\partial m_j} = \frac{s_i(t, m_j^0 + \Delta m_j) - s_i(t, m_j^0)}{\Delta m_j} \,. \tag{A.8}$$

In equation A.8 $s_i(t, m_j^0 + \Delta m_j)$ represents the synthetics simulated by adding to the initial source parameter m_j^0 a suitable perturbation Δm_j.

Possible conditions can be applied to the inversion. For example, imposing a zero-trace constraint, one requires that the trace of the moment tensor matrix $\mathbf{M}$ is zero (no change in volume). Additionally, one can require that the source is represented by a double-couple mechanism by applying a double-couple constraint, i.e. $\det(\mathbf{M}) = 0$.

A.2 A preliminary test

Liu et al. (119) (2004)'s source inversion is based upon the theory presented in section A.1 and implemented in the code CMT3D. However, we also need to simulate the synthetics for the initial and new models and the synthetics required to construct the Fréchet derivatives A.7 - A.8. To this end we use the spectral-element code SPECFEM3D 2.0. Moreover, `FLEXWIN` is required in order to select the windows of the observed and synthetic timeseries that will be involved in the waveform misfit function A.1. The main steps of the procedure for each event are summarized in what follows.

1. Simulate the synthetics for the initial source model $\mathbf{m}_s^0$ and a given velocity model $\mathbf{m}_{vel}$ (1 SEM simulation).

2. Simulate the synthetics $s_i(t, M_j^0)$ for each of the initial moment tensor components (6 SEM simulations).

3. If one wants to invert also for event location (depth, latitude, longitude), simulate the synthetics $s_i(t, m_j^0 + \Delta m_j)$ for each parameter (3 SEM simulations). Δm_j should be chosen carefully to avoid numerical noise and to effectively perturb the source position.

4. Apply the same processing to the data and to all the synthetics.

5. Select the measurement windows by comparing the data and the synthetics for the initial source model $\mathbf{m}_s^0$ and by using `FLEXWIN`. Only these windows are considered in the inversion.

6. Use the code CMT3D. It calculates the Fréchet derivatives in A.7 - A.8 and solves the inverse problem A.4 determining the new source model $\mathbf{m}_s^{new}$ that minimizes the misfit function A.1. A solution can be determined for both zero-trace and double-couple constraints. The quality of each new source model is assessed by estimating its misfit reduction with respect to the initial model.

7. Simulate the synthetics for the new source model $\mathbf{m}_s^{new}$ and the same velocity model $\mathbf{m}_{vel}$ (1 SEM simulation), and process them.

We already ran the SEM simulations for the initial model (step 1; section 4.5.1). Therefore, to invert for both moment tensor and location, for each event one would need to perform 9 additional SEM simulations, plus another simulation to obtain the new synthetics (step 7). Based upon section 4.10, considering timeseries of 60 s and using 256 cores of the INGV clusters, this would require ∼6.7 hours for each event, i.e. a total of 420 hours for all the 63 events (∼1.1x10^5 CPU hours). Due to our limited computational time, these CMT inversions are too expensive to be included in our thesis work together with the adjoint tomography. However, for the future developments, besides the INGV HPC resources, we presently gained additional 2x10^4 CPU hours at CINECA HPC center, that will help to manage the simulation requirements.

A very preliminary test has been performed before starting our tomographic inversion, to have a first insight on how the CMT inversion proceeds.

We considered the M_w 4.8 April 6th 2009 event occurred at 02:37 UTC and we followed the above described steps. The initial source model is given by the TDMT solution and we used our initial 3D velocity model for central Italy (section 4.2.1), now labeled $\mathbf{m}^0_{vel}$. The data from the velocimetric stations that recorded this event (section 4.4), together with all the synthetics, have been processed based on the procedure in section 4.5.1. The filtering period band is 2-20 s and `FLEXWIN` selected the windows within this range. Then, CMT3D was used to invert for the 6 moment tensor components of the event, the MT components and the depth (7 parameters), and the MT components, the depth and the horizontal location (9 parameters). Each of these solutions was constrained by either a zero-trace or also a double-couple condition, for a total of six possible new CMT solutions for the considered event.

The results obtained from CMT3D are presented in Figure A.1 (a) and Table A.1. One notes that, in general, the new source mechanisms are consistent with the TDMT solution and the updated location parameters (when inverted) are only slightly different from the original CMT. On the contrary, the resulting seismic moment in all cases is significantly reduced compared to the initial M_0 (up to $\sim$45%). The solutions with the same constraint (zero-trace 'zt' or double-couple 'dc') are generally similar to each other, also considering their misfit reductions ($[(F(\mathbf{m}^0_s) - F(\mathbf{m}^{new}_s))/F(\mathbf{m}^0_s)] * 100$).

To choose the result that seems more suitable to be the new CMT solution, we consider that adding the double-couple constraint usually provides more stable solutions. Moreover, the misfit reduction is taken into account and the solutions that do not differ too much from the original CMT are preferred. Considering all these aspects, sometimes in contrast to each other, we finally selected the solution obtained by inverting for 7 parameters with also a double-couple constraint (7par-dc). The new synthetics (blu) are compared to the original ones (red) and to the data (black) in Figure A.1 (b). The synthetics for the inverted solution have generally lower amplitudes with respect to the original synthetics, as expected from the seismic moment reduction. The waveform changes, but the fit of the observed timeseries is not always improved.

These are very preliminary results. They seem to require a reduction of the seismic moment of the event, that could be consistent with some considerations of our work (chapter 3). In fact, combining a TDMT source solution, based on a 1D velocity profile,

with a 3D structure model, may result in an implicit incompatibility, that probably produces overestimated amplitudes. However, this is only one example and the reliability of the results necessarily requires to be assessed with further analyses. In particular, one should remark that TDMT solutions are obtained using recordings in the period band 10-50 s, that differs from the range that we consider (2-20 s). Thus, a more careful comparison in the same period range is required. Moreover, the variations of the event location parameters (depth, latitude, longitude) seem too small to be reliable. Therefore, further tests are also needed with more suitable model perturbations[1]. Note that the choice of the new solution is somehow subjective and the one with the highest misfit reduction is not necessarily the best CMT solution.

Although accounting for the discussed issues, this preliminary test seems to suggest that a refinement of the source parameters produces variations in waveform shape and amplitudes that could help to fit the data, also using our initial velocity model $\mathbf{m}^0_{vel}$. We think that trying this CMT inversion with our updated model $\mathbf{m}^5_{vel}$ may significantly contribute to reduce the misfit between data and synthetics.

	Misf.Red. [%]	M_0 [10^{23}dyne·cm]	Δdepth [10^{-3}km]	Δlat [m]	Δlon [m]
6par-zt	62.6	1.47			
6par-dc	59.0	1.29			
7par-zt	62.6	1.47	-0.920		
7par-dc	58.8	1.31	0.198		
9par-zt	62.7	1.47	-20.0	75.3	-21.2
9par-dc	58.8	1.27	-206.0	57.1	-240.0

Table A.1: Example of 3D CMT inversion - For each solution, provided by CMT3D for the 6/4/2009 02:37 UTC event, the table reports the misfit reduction (Misf.Red.=$[(F(\mathbf{m}^0_s) - F(\mathbf{m}^{new}_s))/F(\mathbf{m}^0_s)] * 100$), the new seismic moment M_0, the depth and horizontal coordinate variations with respect to the initial source parameters ($\Delta m_j = m^{new}_j - m^0_j$). See caption of Figure A.1 (a) for the label of each solution.

[1]Presently we use a Δm_j=1 km in A.8 for all the 3 parameters. It coincides with the minimum size of the mesh elements (section 4.2.2), but probably does not perturb enough the source location.

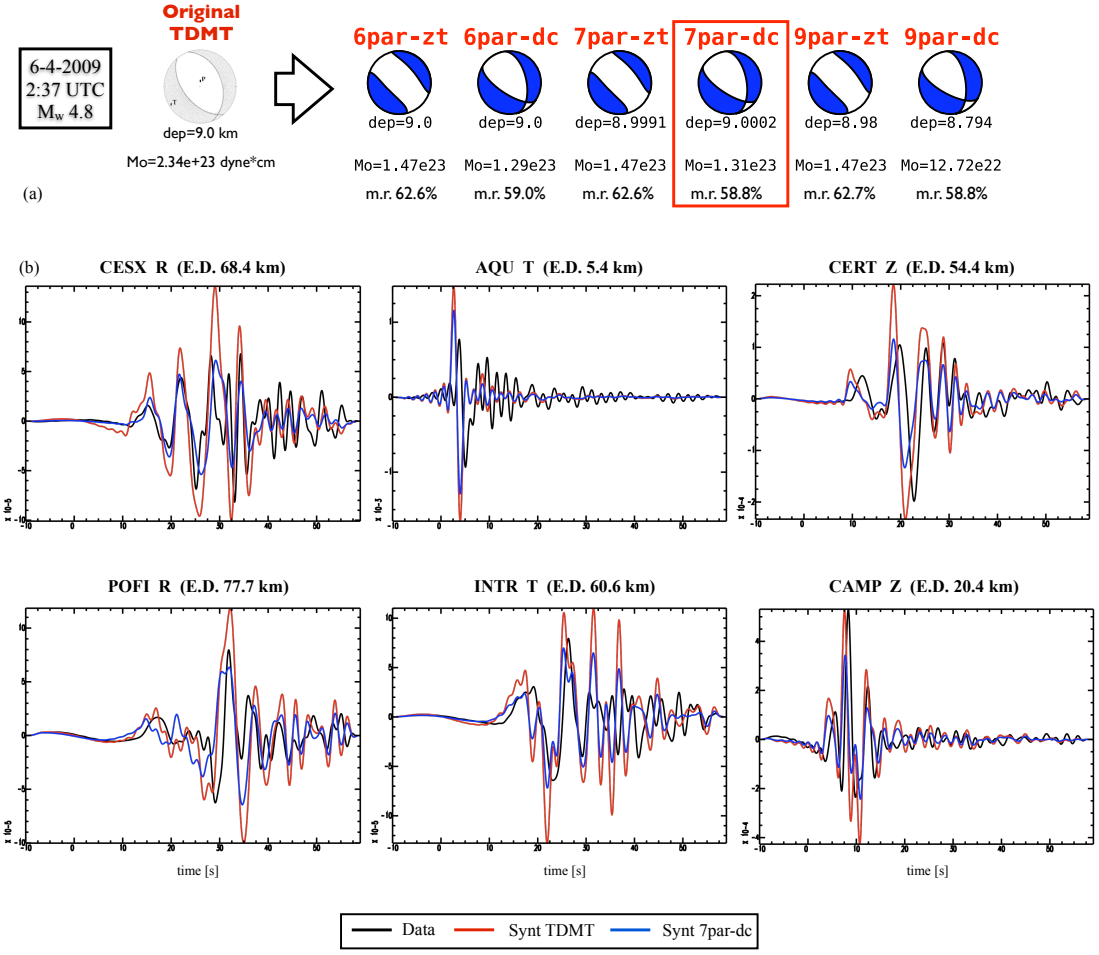


Figure A.1: Example of 3D CMT inversion - (a) For the given event, the original TDMT represents the initial Time Domain Moment Tensor solution. The six solutions on the right of the arrow are the results obtained by using CMT3D and inverting for the 6 moment tensor components with zero-trace or also double-couple constraints (6par-zt or 6par-dc, respectively), for the MT components and the depth with the two possible constraints (7par-zt or 7par-dc), and for the MT components, the depth and the horizontal location in the two cases (9par-zt or 9par-dc). For each new source model $\mathbf{m}_s^{new}$ we indicate the new depth and scalar seismic moment and the misfit reduction with respect to the initial source model $\mathbf{m}_s^0$ (m.r.=$[(F(\mathbf{m}_s^0) - F(\mathbf{m}_s^{new}))/F(\mathbf{m}_s^0)] * 100$). See also Tab. A.1. (b) Comparisons of the data (black), the synthetics for the original TDMT (red), and the synthetics for the chosen inverted solution (blue), i.e. 7par-dc (red square in fig.(a)). The name of the station, its epicentral distance (E.D.) and the considered component are indicated for each seismogram. The timeseries are displacements (in m) filtered between 2-20 s.

Appendix B

FLEXWIN user parameters

Standard tuning parameters:	
$T_{0,1}$	bandpass filter corner periods
$r_{P,A}$	signal to noise ratios for whole waveform
$r_0(t)$	signal to noise ratios single windows
$w_E(t)$	water level on short-term/long-term ratio
$CC_0(t)$	acceptance level for normalized cross-correlation
$\Delta\tau_0(t)$	acceptance level for time lag
$\Delta\ln A_0(t)$	acceptance level for amplitude ratio
$\Delta\tau_{ref}$	reference time lag
$\Delta\ln A_{ref}$	reference amplitude ratio
Fine tuning parameters:	
c_0	for rejection of internal minima
c_1	for rejection of short windows
c_2	for rejection of un-prominent windows
$c_{3_{a,b}}$	for rejection of multiple distinct arrivals
$c_{4_{a,b}}$	for curtailing of windows with emergent starts and/or codas
w_{CC} w_{len} w_{nwin}	for selection of best non-overlapping window combination

Table B.1: Overview of standard tuning parameters and of fine tuning parameters. Values are defined in a parameter file, and the time dependence of those that depend on time is described by user-defined functions. Courtesy of Maggi et al. (125) (2009).

In section 4.5.2 we outlined the steps of the time window selection procedure based upon the code `FLEXWIN` (Maggi et al. (125), 2009). The algorithm involves some user

parameters, that define the conditions to establish if a window should be retained or rejected. They need to be tuned depending on the considered seismological scenario and are set in a parameter file input of the code. In Table B.1, we summarize the description of these parameters and in Table B.2 the values chosen for the tomography of central Italy.

Note that some of the parameters are time dependent. Thus, the user can describe their behavior in time by defining their functional form in the code, in order to modify the applied conditions along the time series. In our case, the time-dependence of the water level $w_E(t)$ is chosen so that the algorithm could select windows containing either body and surface waves.

For more details on the procedure and the user parameters refer to Maggi et al. (125) (2009).

$T_{0,1}$	2-20 s	6-20 s
$r_{P,A}$	2.5; 3.5	3.0; 2.5
$r_0(t)$	4.0	3.0
$w_E(t)$	0.07	0.08
CC_0	0.75	0.71
$\Delta\tau_0$	3.0	8.0
$\Delta \ln A_0$	1.5	1.5
$\Delta\tau_{ref}$	1.0	4.0
$\Delta \ln A_{ref}$	0.0	0.0
c_0	1.0	0.7
c_1	5.0	2.0
c_2	0.0	0.0
$c_{3_{a,b}}$	4.0; 2.5	3.0; 2.0
$c_{4_{a,b}}$	2.0; 6.0	2.5; 12.0
w_{CC} w_{len} w_{nwin}	1.0; 1.0; 1.0	1.0; 1.0; 1.0

Table B.2: `FLEXWIN` **user parameters for central Italy tomographic inversion**. The values are assigned by performing some window selection tests for the considered scenario and period ranges. Only the water level $w_E(t)$ is chosen to be time-dependent and its functional form allows us to select both body and surface wave windows.

Appendix C

`measure_adj` user parameters

The estimate of the measurements, the misfit functions and the adjoint sources needed for the inversion has been performed by using the code `measure_adj` (section 4.6.1). The program receives in input the windows selected by the code `FLEXWIN` (section 4.5.2) and it performs the calculations accounting for some user parameters. One of the operations is to establish if a multitaper measurement made in a given window is reasonable enough to be retained, or it should be reverted to a cross-correlation measurement. This assessment is based upon the multitaper parameters reported in Table C.1, that should be chosen depending on the considered dataset. Table C.1 also reports the values chosen for the study of central Italy tomography. For more details refer to the code `measure_adj` and to its user manual [1].

Ratio between current period and timestep of SEM simulations:	DT_FAC = 2.0
Ratio between current period and error of traveltime estimate:	ERR_FAC = 2.5
Max time shift allowed at all frequency should be (DT_MAX_SCALE * Tshift_CC) DT_MAX_SCALE = 3.5	
Number of cycles in a window:	NCYCLE_IN_WINDOW = 1.5

Table C.1: `measure_adj` **user parameters for central Italy tomographic inversion**. If conditions on these chosen user parameters are satisfied by the considered time window (see the code[1]), the corresponding multitaper measurement is retained; otherwise it is reverted to a cross-correlation (CC) measurement. The timestep of SEM simulations is fixed at $1x10^{-3}$ s; the error of traveltime (TT) estimate and the CC-TT measurement Tshift_CC are estimated by `measure_adj` for the given window.

[1] http://www.geodynamics.org/svn/cig/seismo/3D/ADJOINT_TOMO/measure_adj/

Acknowledgements

During these three years many things changed in my life. Working at INGV offered me the opportunity to know very interesting people and new friends, and to taste the seismological research world. The friendly and informal environment of the institute favors collaboration and enriched my work with new ideas and perspectives.

I can't forget how my tutors Alberto and Emanuele welcomed me. I was immediately involved in challenging and exciting projects, that greatly contributed to my scientific and personal background. They were always willing to answer my questions, and our discussions encourage my scientific curiosity and were essential for my work. They also took care of my computational education with everyday new precious teachings.

I am grateful to Alberto, Emanuele and Manuela who shared their offices with me, as soon as I came to INGV and then for the whole duration of my PhD. They were and are officemates, colleagues and friends, who aided me in every situation and from whom I learned a lot. I thank all the other researchers I worked with and the ones who helped me to manage bureaucratic matters. I also appreciated the strong support from all the people of the INGV HPC, these resources were precious for my thesis. In particular, Daniele was always willing to help me both for 'cluster' matters and in general for other unexpected issues - included a sudden undesired 'summertime holiday' of my computer when I was working at my thesis!

The PhD provides me the wonderful opportunity of living and studying abroad for six months. I thank Prof. Jeroen Tromp for its hospitality and teachings, that were fundamental for my work. In Princeton I met friends and colleagues who shared with me this experience, contributing to make it great.

My friends, the near and the distant ones, they all participated in this exciting path, encouraging me at every personal and professional test and sharing in turn their experiences with me. My family supported me - and stood me! - in all the situations and difficulties. Any thanks is not enough for my parents, who once again were next to me, always helping to find the good side of things and enjoy the life. Paolo deserves an extraordinary thanks for giving me everyday the strength and enthusiasm for never give up.

Finally, I wish to dedicate this thesis to my grandmother and my uncle. They lived differently but both intensely, they always believed in me and they will be always part of my life.

References

[1] Akçelik, V., Biros, G., & Ghattas, O., 2002. Parallel multiscale Gauss–Newton–Krylov methods for inverse wave propagation, *Proc. ACM/IEEE Supercomputing SC'2002 conference*, published on CD-ROM and at www.sc-conference.org/sc2002. 8, 49

[2] Akçelik, V., Bielak, J., Biros, G., Epanomeritakis, I., Fernández, A., Ghattas, O., Kim, E. J., López, J., O'Hallaron, D., Tu, T., & Urbanic, J., 2003. High resolution forward and inverse earthquake modeling on terascale computers, *Proc. ACM/IEEE Supercomputing SC'2003 conference*, published on CD-ROM and at www.sc-conference.org/sc2003. 8, 49

[3] Aki, K. & Richards, P. G., 1980. *Quantitative seismology, theory and methods*, W. H. Freeman, San Francisco, USA. 1, 29, 74

[4] Akinci, A., Galadini, F., Pantosti, D., Petersen, M., Malagnini, L., & Perkins, D., 2009. Effect of time dependence on probabilistic seismic-hazard maps and deaggregation for the Central Apennines, Italy, *Bull. Seismol. Soc. Am.*, **99**(2A), 585–610. 99

[5] Akinci, A., Malagnini, L., & Sabetta, F., 2010. Characteristics of the strong ground motions from the 6 April 2009 LAquila earthquake, Italy, *Soil Dynam. Earth. Engin.*, **30**, 320–335. 60, 63, 64, 66, 85, 88, 99

[6] Alessandrini, B., Beranzoli, L., & Mele, F., 1995. 3-D crustal P-wave velocity tomography of the Italian region using local and regional seismicity data, *annali di geofisica*, **38**(2), 189–211. 60

[7] Amato, A. & Mele, F., 2008. Performance of the INGV National Seismic Network from 1997 to 2007, *Annals of Geophysics*, **51**, 417–431. 65

[8] Ameri, G., Massa, M. P., Bindi, D., DAlema, E., Gorini, A., Luzi, L., Marzorati, S., Pacor, F., Paolucci, R., Pugli, R., & Smerzini, C., 2009. The April 6 2009, Mw 6.3 LAquila (central Italy) earthquake: strong-motion observations, *Seismol. Res. Lett.*, **80**, 951–966. 88

[9] Anzidei, M., Boschi, E., Cannelli, V., Devoti, R., Esposito, A., Galvani, A., Melini, D., Pietrantonio, G., Riguzzi, F., Sepe, V., & Serpelloni, E., 2009. Coseismic deformation of the destructive April 6, 2009 L'Aquila earthquake (central Italy) from GPS data, *Geophys. Res. Lett.*, **36**, L17307. 63

[10] Atzori, S., Hunstad, I., Chini, M., Salvi, S., Tolomei, C., Bignami, C., Stramondo, S., Trasatti, E., Antonioli, A., & Boschi, E., 2009. Finite fault inversion of DInSAR coseismic displacement of the 2009 LAquila earthquake (central Italy), *Geophys. Res. Lett.*, **36**, L15305. 60, 63

[11] Bagh, S., Chiaraluce, L., De Gori, P., Moretti, M., Govoni, A., Chiarabba, C., Di Bartolomeo, P., & Romanelli, M., 2007. Background seismicity in the central Apennines of Italy: The Abruzzo region case study, *Tectonophysics*, **444**(1-4), 80–92. 67, 68

[12] Bao, H., Bielak, J., Ghattas, O., Kallivokas, L. F., O'Hallaron, D. R., Shewchuk, J. R., & Xu, J., 1998. Large-scale simulation of elastic wave propagation in heterogeneous media on parallel computers, *Comput. Meth. Appl. Mech. Eng.*, **152**, 85–102. 3

[13] Bérenger, J. P., 1994. A Perfectly Matched Layer for the absorption of electromagnetic waves, *J. Comput. Phys.*, **114**, 185–200. 14

[14] Bijwaard, H. & Spakman, W., 2000. Nonlinear global P-wave tomography by iterated linearized inversion, *Geophys. J. Int.*, **141**, 71–82. 6

[15] Bindi, D., Castro, R., Franceschina, G., Luzi, L., & Pacor, F., 2004. The 1997–1998 Umbria-Marche sequence (central Italy): source, path, and site effects estimated from strong motion data recorded in the epicentral area, *J. Geophys. Res.*, **109**(B4), B04312. 88, 99

[16] Bindi, D., Pacor, F., Luzi, L., Massa, M., & Ameri, G., 2009. The Mw 6.3, 2009 L'Aquila earthquake: source, path and site effects from spectral analysis of strong motion data, *Geophys. J. Int.*, pp. 1–7. 66, 88, 103

[17] Blacker, T., Knupp, P., Lober, R. R., Melander, D., Sjaardema, G. D., Sample, W. A., Smith, M. K., Tautges, T. J., White, D. R., Benzley, S., Kerr, R., Jankovich, S. R., McRae, D. B., & Panthaki, M., 1994. *CUBIT Mesh Generation Environment Users Manual Vol. 1*, Sandia National Laboratories, Albuquerque, NM. 50

[18] Bosi, C. & Bertini, T., 1970. Geologia della media valle dellAterno, *Mem. Soc. Geol. It.*, **IX**, 719–777. 63

[19] Bouchon, M. & Barker, J. S., 1996. Seismic response of a hill: the example of Tarzana, California, *Bull. Seismol. Soc. Am.*, **86**(1A), 66–72. 84

[20] Bouchon, M., Schultz, C. A., & Töksoz, M. N., 1996. Effect of three-dimensional topography on seismic motion, *J. Geophys. Res.*, **101**, 5835–5846. 2, 84

[21] Canuto, C., Hussaini, M. Y., Quarteroni, A., & Zang, T. A., 1988. *Spectral methods in fluid dynamics*, Springer-Verlag, New York, USA. 20, 22

[22] Capdeville, Y., Chaljub, E., Vilotte, J. P., & Montagner, J. P., 2003. Coupling the spectral element method with a modal solution for elastic wave propagation in global Earth models, *Geophys. J. Int.*, **152**, 34–67. 3

[23] Carcione, J. M., 1994. The wave equation in generalized coordinates, *Geophysics*, **59**, 1911–1919. 2, 26

[24] Casarotti, E., Stupazzini, M., Lee, S., Komatitsch, D., Piersanti, A., & Tromp, J., 2008. CUBIT and seismic wave propagation based upon the spectral-element method: An advanced unstructured mesher for complex 3D geological media, in *Proceedings of the 16th International Meshing Roundtable*, vol. 5B.4, pp. 579–597. 16, 19, 51

[25] Castro, R., Pacor, F., Bindi, D., Franceschina, G., & Luzi, L., 2004. Site response of strong motion stations in the Umbria, central Italy, region, *Bull. Seismol. Soc. Am.*, **94**, 576–590. 88, 103

[26] Cavinato, G. & De Celles, P., 1999. Extensional basins in the tectonically bimodal central Apennines fold-thrust belt, Italy: response to corner flow above subducting slab retrograde motion, *Geology*, **27**, 955–958. 63

[27] Çelebi, M., 1987. Topographical and geological amplifications determined from strong-motion and aftershock records of the 3 March 1985 Chile earthquake, *Bull. Seismol. Soc. Am.*, **77**, 1147–1167. 84

[28] Çelebi, M., Bazzurro, P., Chiaraluce, L., Clemente, P., Decanini, L., DeSortis, A., Ellsworth, W., Gorini, A., Kalkan, E., Marcucci, S., Milana, G., Mollaioli, F., Olivieri, M., Paolucci, R., Rinaldis, D., Rovelli, A., Sabetta, F., & Stephens, C., 2010. Recorded motions of the 6 April 2009 Mw 6.3 LAquila, Italy, earthquake and implications for building structural damage: Overview, *Earthquake Spectra*, **26**, 651–684. 76, 88

[29] Chaljub, E., 2000. *Modélisation numérique de la propagation d'ondes sismiques en géométrie sphérique : application à la sismologie globale (Numerical modeling of the propagation of seismic waves in spherical geometry: application to global seismology)*, Ph.D. thesis, Université Paris VII Denis Diderot, Paris, France. 3

[30] Chaljub, E. & Valette, B., 2004. Spectral element modelling of three-dimensional wave propagation in a self-gravitating Earth with an arbitrarily stratified outer core, *Geophys. J. Int.*, **158**, 131–141. 3

[31] Chaljub, E., Cornou, C., Guéguen, P., Causse, M., & Komatitsch, D., 2005. Spectral-element modeling of 3D wave propagation in the alpine valley of Grenoble, France, in *Geophysical Research Abstracts*, vol. 7, 05225, EGU 2nd general assembly, Wien, Austria. 4

[32] Chaljub, E., Komatitsch, D., Vilotte, J. P., Capdeville, Y., Valette, B., & Festa, G., 2007. Spectral element analysis in seismology, in *Advances in wave propagation in heterogeneous media*, vol. 48 of **Advances in Geophysics**,

pp. 365–419, eds Wu, R.-S. & Maupin, V., Elsevier - Academic Press, London, UK. 2, 4, 14, 88

[33] Chaljub, E., Moczo, P., Tsuno, S., Bard, P.-Y., Kristek, J., Käser, M., Stupazzini, M., & Kristekova, M., 2010. Quantitative comparison of four numerical predictions of 3D ground motion in the Grenoble valley, France, *Bull. Seismol. Soc. Am.*, **100**(4), 1427–1455. 3, 4

[34] Cheloni, D., D'Agostino, N., D'Anastasio, E., Avallone, A., Mantenuto, S., Giuliani, R., Mattone, M., Calcaterra, S., Gambino, P., & Dominici, D., 2010. Coseismic and initial postseismic slip of the 2009 Mw 6.3 L'Aquila earthquake, Italy, from GPS measurements, *Geophys. J. Int.*, **181**(3), 1539–1546. 60

[35] Chen, M. & Tromp, J., 2007. Theoretical and numerical investigations of global and regional seismic wave propagation in weakly anisotropic earth models, *Geophys. J. Int.*, **168**(3), 1130–1152. 54

[36] Chen, P., Zhao, L., & Jordan, T. H., 2007. Full 3D tomography for the crustal structure of the Los Angeles region, *Bull. Seismol. Soc. Am.*, **97**(4), 1094–1120. 191

[37] Chevalier, C. & Pellegrini, F., 2008. PT-SCOTCH: A tool for efficient parallel graph ordering, *Parallel Computing*, **34**(6-8), 318–331. 24, 51

[38] Chiarabba, C. & Amato, A., 1996. Crustal velocity structure of the Apennines (Italy) from P-wave travel time tomography, *Annali di Geofisica*, **39**, 1133–1148. 60

[39] Chiarabba, C., Bagh, S., Bianchi, I., De Gori, P., & Barchi, M., 2010. Deep structural heterogeneities and the tectonic evolution of the Abruzzi region (Central Apennines, italy) revealed by microseismicity, seismic tomography, and teleseismic receiver functions. Doi:10.1016/j.epsl.2010.04.028, *Earth Planet. Sci. Lett.*. 60, 63, 69, 71, 94, 96, 159, 160, 161

[40] Chiaraluce, L., Amato, A., Cocco, M., Chiarabba, C., Selvaggi, G., Bona, M. D., Piccinini, D., Deschamps, A., Margheriti, L., Courboulex, F., & Ripepe, M., 2004. Complex normal faulting in the Apennines thrust-and-fold belt: The 1997 seismic sequence in central Italy, *Bull. Seismol. Soc. Am.*, **94**, 99–116. 63

[41] Cirella, A., Piatanesi, A., Tinti, E., & Cocco, M., 2006. Dynamically consistent source time functions to invert kinematic rupture histories, in *Abstract S41B-1323 presented at 2006 AGU Fall Meeting*, San Francisco, California, USA. 80, 82, 83

[42] Cirella, A., Piatanesi, A., Cocco, M., Tinti, E., Scognamiglio, L., Michelini, A., Lomax, A., & Boschi, E., 2009. Rupture history of the 2009 L'Aquila (Italy) earthquake from non-linear joint inversion of strong motion and GPS data, *Geophys. Res. Lett.*, **36**, L19304. 60, 63, 64, 66, 67, 68, 69, 73, 74, 75, 76, 78, 79, 80, 81, 82, 83, 85, 94, 96, 99, 102, 103, 104, 108, 194, 199

[43] Cirella, A., Piatanesi, A., Tinti, E., Chini, M., & Cocco, M., 2011. Complexity of the rupture process during the 2009 LAquila, Italy, earthquake, *Submitted to Geophys. J. Int.*. 104

[44] Clayton, R. & Engquist, B., 1977. Absorbing boundary conditions for acoustic and elastic wave equations, *Bull. Seismol. Soc. Am.*, **67**, 1529–1540. 14, 54

[45] Clouser, R. H. & Langston, C. A., 1995. Modeling observed P-R_g conversions from isolated topographic features near the NORESS array, *Bull. Seismol. Soc. Am.*, **85**, 859–873. 85

[46] Cohen, G., Joly, P., & Tordjman, N., 1993. Construction and analysis of higher-order finite elements with mass lumping for the wave equation, in *Proceedings of the second international conference on mathematical and numerical aspects of wave propagation*, pp. 152–160, SIAM, Philadelphia, Pennsylvania, USA. 3

[47] Courant, R., Friedrichs, K. O., & Lewy, H., 1928. Über die partiellen Differenzengleichungen der mathematischen Physik, *Mathematische Annalen*, **100**, 32–74. 25

[48] Crase, E., Pica, A., Noble, M., McDonald, J., & Tarantola, A., 1990. Robust elastic non-linear waveform inversion: application to real data, *Geophysics*, **55**, 527–538. 7, 28

[49] Dahlen, F., S.-H., H., & G., N., 2000. Fréchet kernels for finite-frequency traveltimes-I. Theory, *Geophys. J. Int.*, **141**, 157–174. 4, 34, 35

[50] Dahlen, F. A. & Baig, A. M., 2002. Fréchet kernels for body-wave amplitudes, *Geophys. J. Int.*, **150**, 440–466. 5

[51] Dahlen, F. A. & Tromp, J., 1998. *Theoretical Global Seismology*, Princeton University Press, Princeton, New-Jersey, USA. 1, 13, 14, 29, 116, 117

[52] de la Puente, J., Käser, M., Dumbser, M., & Igel, H., 2007. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes IV: Anisotropy, *Geophys. J. Int.*, **169**, 1210–1228. 3

[53] de la Puente, J., Dumbser, M., Käser, M., & Igel, H., 2008. Discontinuous Galerkin methods for wave propagation in poroelastic media, *Geophysics*, **73**, T77–T97. 3

[54] De Luca, G., Marcucci, S., Milana, G., & Sano, T., 2005. Evidence of low-frequency amplification in the city of L'Aquila, central italy, through a multidisciplinary approach including strong-and weak-motion data, ambient noise, and numerical modeling, *Bull. Seismol. Soc. Am.*, **95**(4), 1469–1481. 63, 88

[55] Delavaud, E., Cupillard, P., Festa, G., & Vilotte, J.-P., 2006. 3D spectral element method simulations of the seismic response in the Caracas basin, in *Proc. of the Third International Symposium on the Effects of Surface Geology on Seismic Motion*, vol. 1, pp. 512–522, Grenoble, France. 4

[56] Di Luccio, F., Ventura, G., Di Giovambattista, R., Piscini, A., & Cinti, F. R., 2010. Normal faults and thrusts re-activated by deep fluids: the 6 April 2009 Mw 6.3 LAquila earthquake, central Italy, *J. Geophys. Res.*, **115**, B06315. 162, 188

[57] Di Luzio, E., Mele, G., Tiberti, M. M., Cavinato, G. P., & Parotto, M., 2009. Moho deepening and shallow upper crustal delamination beneath the central Apennines, *Earth Planet. Sci. Lett.*, **280**, 1–12. 70

[58] Di Stefano, R., Chiarabba, C., & Lucente, F., 1999. Crustal and uppermost mantle structure in Italy from the inversion of P-wave arrival times: Geodynamic implications, *Geophys. J. Int.*, **139**, 483–498. 60

[59] Di Stefano, R., Kissling, E., Chiarabba, C., Amato, A., & Giardini, D., 2009. Shallow subduction beneath Italy: three-dimensional images of the Adriatic-European-Tyrrhenian lithosphere system based on high-quality P wave arrival times., *J. Geophys. Res.*, **114**, B05305. 60

[60] Di Stefano, R., Bianchi, I., Ciaccio, M. G., Carrara, G., & Kissling, E., 2011. Three-dimensional Moho topography in Italy: New constraints from receiver functions and controlled source seismology, *Geochem. Geophys. Geosyst.*, **12**, Q09006, doi:10.1029/2011GC003649. 60, 70, 71, 72, 113, 114, 115

[61] Doglioni, C., 1991. A proposal of kinematic modelling for W?dipping subductions - possible applications to the Tyrrhenian?Apennines system, *Terra Nova*, **3**, 423–434. 62

[62] Doglioni, C., 1995. Geological remarks on the relationships between extension and convergent geodynamic settings, *Tectonophysics*, **252**, 253–267. 62

[63] Donati, S., Marra, F., & Rovelli, A., 2001. Damage and ground shaking in the town of Nocera Umbra during Umbria-Marche, central Italy, earthquakes: The special effect of a fault zone, *Bull. Seismol. Soc. Am.*, **91**, 511–519. 84

[64] Dreger, D. S. & Helmberger, D. V., 1993. Determination of source parameters at regional distances with 3-component sparse network data, *J. Geophys. Res.*, **98**, 8107–8125. 115

[65] Dumbser, M. & Käser, M., 2006. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes-II. The three-dimensional isotropic case, *Geophys. J. Int.*, **167**(1), 319–336. 3

[66] Dziewonski, A., Chou, T.-A., & Woodhouse, J., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.*, **86**, 2825–2852. 116

[67] Ekström, G., Tromp, J., & Larson, E. W. F., 1997. Measurements and global models of surface wave propagation, *J. Geophys. Res.*, **102**, 8137–8157. 123

[68] EMERGEO, Working Group 2010. Evidence for surface rupture associated with the M_w 6.3 L'Aquila earthquake sequence of April 2009 (central Italy), *Terra Nova*, **22**, 43–51. 63

[69] Emmerich, H. & Korn, M., 1987. Incorporation of attenuation into time-domain computations of seismic wave fields, *Geophysics*, **52**, 1252–1264. 12

[70] Faccioli, E., Maggio, F., Paolucci, R., & Quarteroni, A., 1997. 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method, *J. Seismol.*, **1**, 237–251. 3

[71] Faenza, L., Lauciani, V., & Michelini, A., 2011. Rapid determination of the shakemaps for the L'Aquila main shock: a critical analysis, *Boll. Geof. Teor. Appl.*, **52**(3), 1–19. 100, 101

[72] Fichtner, A., Bunge, H.-P., & Igel, H., 2006a. The adjoint method in seismology - I. Theory, *Phys. Earth Planet. In.*, **157**(1-2), 86–104. 6

[73] Fichtner, A., Bunge, H.-P., & Igel, H., 2006b. The adjoint method in seismology - II. Applications: traveltimes and sensitivity functionals, *Phys. Earth Planet. In.*, **157**(1-2), 105–123. 6

[74] Fichtner, A., Igel, H., Bunge, H.-P., & Kennett, B. L. N., 2009a. Simulation and inversion of seismic wave propagation on continental scales based on a spectral-element method, *Journal of Numerical Analysis, Industrial and Applied Mathematics*, **4**(1-2), 11–22. 3

[75] Fichtner, A., Kennett, B. L. N., Igel, H., & Bunge, H.-P., 2009b. Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophys. J. Int.*, **179**, 1703–1725. 8, 190

[76] Fletcher, R. & Reeves, C., 1964. Function minimization by conjugate gradients, *Comp. J.*, **7**, 149–154. 7

[77] Furumura, T., Kennett, B. L. N., & Furumura, M., 1998. Seismic wavefield calculation for laterally heterogeneous whole Earth models using the pseudospectral method, *Geophys. J. Int.*, **135**(3), 845–860. 2

[78] Gauthier, O., Virieux, J., & Tarantola, A., 1986. Two-dimensional non-linear inversion of seismic waveforms: numerical results, *Geophysics*, **51**, 1387–1403. 7, 55

[79] Graves, R. W., 1996. Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences, *Bull. Seismol. Soc. Am.*, **86**(4), 1091–1106. 2

[80] Gropp, W., Lusk, E., & Skjellum, A., 1994. *Using MPI, portable parallel programming with the Message-Passing Interface*, MIT Press, Cambridge, USA. 24

[81] Herrmann, B. & Malagnini, L., 2009. Systematic determination of moment tensor of the April 6th, 2009 LAquila earthquake sequence, in *Abstract U23A-0029 presented at 2009 AGU Fall Meeting*, vol. 90, San Francisco, California, USA. 67

[82] Hudson, J. A., 1977. Scattered waves in the coda of p, *J. Geophys.*, **43**, 359–374. 29

[83] Hughes, T. J. R., 1987. *The finite element method, linear static and dynamic finite element analysis*, Prentice-Hall International, Englewood Cliffs, New Jersey, USA. 17, 25

[84] Hung, S.-H., Dahlen, F. A., & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes-II. Examples, *Geophys. J. Int.*, **141**, 175–203. 5

[85] Hunstad, I., Selvaggi, G., DAgostino, N., England, P., Clarke, P., & Pierozzi, M., 2003. Geodetic strain in peninsular Italy between 1875 and 2001, *Geophys. Res. Lett.*, **30**, 1181. 62

[86] Husen, S. & Kissling, E., 2001. Postseismic fluid flow after the large subduction earthquake of Antofagasta, Chile, *Geology*, **29**, 847–850. 162

[87] Igel, H. & Weber, M., 1995. SH-wave propagation in the whole mantle using high-order finite differences, *Geophys. Res. Lett.*, **22**, 731–734. 2

[88] Igel, H. & Weber, M., 1996. P-SV wave propagation in the whole mantle using high-order finite differences: application to lowermost mantle structure, *Geophys. Res. Lett.*, **23**, 415–418. 2

[89] Igel, H., Mora, P., & Riollet, B., 1995. Anisotropic wave propagation through finite-difference grids, *Geophysics*, **60**(4), 1203–1216. 2

[90] Jarvis, A., Reuter, H., Nelson, A., & Guevara, E., 2008. Hole-filled seamless SRTM data V4. 71, 114

[91] Karypis, G. & Kumar, V., 1998. A parallel algorithm for multilevel graph partitioning and sparse matrix ordering, *Journal of Parallel and Distributed Computing*, **48**, 71–85. 51

[92] Käser, M. & Dumbser, M., 2006. An arbitrary high order discontinuous Galerkin method for elastic waves on unstructured meshes I: The two dimensional isotropic case with external source terms, *Geophys. J. Int.*, **166**, 855–877. 3

[93] Käser, M., Dumbser, M., de la Puente, J., & Igel, H., 2007. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes III: Viscoelastic attenuation, *Geophys. J. Int.*, **168**, 224–242. 3

[94] Kawase, H., 1988. Time-domain response of a semi-circular canyon for incident *SV*, *P* and Rayleigh waves calculated by the discrete wavenumber boundary element method, *Bull. Seismol. Soc. Am.*, **78**, 1415–1437. 2

[95] Kawase, H. & Aki, K., 1990. Topography effect at the critical SV-wave incidence: possible explanation of damage pattern by the Whittier Narrows, California, earthquake of 1 October 1987, *Bull. Seismol. Soc. Am.*, **80**, 1–22. 84

[96] Kennett, B. L. N., 1983. *Seismic wave propagation in stratified media*, Cambridge University Press, Cambridge. 1

[97] Kim, Y., Liu, Q., & Tromp, J., 2011. Adjoint centroid-moment tensor inversions, *Geophys. J. Int.*, **186**(1), 264–278. 4, 8, 9, 74, 104, 108, 199

[98] Komatitsch, D., 1997. *Méthodes spectrales et éléments spectraux pour l'équation de l'élastodynamique 2D et 3D en milieu hétérogène (Spectral and spectral-element methods for the 2D and 3D elastodynamics equations in heterogeneous media)*, Ph.D. thesis, Institut de Physique du Globe, Paris, France, 187 pages. 3

[99] Komatitsch, D. & Martin, R., 2007. An unsplit convolutional Perfectly Matched Layer improved at grazing incidence for the seismic wave equation, *Geophysics*, **72**(5), SM155–SM167. 14

[100] Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral-element method for 3-D seismic wave propagation, *Geophys. J. Int.*, **139**(3), 806–822. 3, 4, 15, 16, 19, 23, 24, 25

[101] Komatitsch, D. & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation-I. Validation, *Geophys. J. Int.*, **149**(2), 390–412. 3, 4, 15

[102] Komatitsch, D. & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation-II. 3-D models, oceans, rotation, and self-gravitation, *Geophys. J. Int.*, **150**(1), 303–318. 3, 4

[103] Komatitsch, D. & Tromp, J., 2003. A Perfectly Matched Layer absorbing boundary condition for the second-order seismic wave equation, *Geophys. J. Int.*, **154**(1), 146–153. 14

[104] Komatitsch, D. & Vilotte, J. P., 1998. The spectral-element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, *Bull. Seismol. Soc. Am.*, **88**(2), 368–392. 3, 4, 15, 20, 25, 84, 85, 98

[105] Komatitsch, D., Vilotte, J. P., Vai, R., Castillo-Covarrubias, J. M., & Sánchez-Sesma, F. J., 1999. The spectral-element method for elastic wave equations: application to 2D and 3D seismic problems, *Int. J. Numer. Meth. Eng.*, **45**(9), 1139–1164. 15

[106] Komatitsch, D., Barnes, C., & Tromp, J., 2000a. Wave propagation near a fluid-solid interface: a spectral element approach, *Geophysics*, **65**(2), 623–631. 4

[107] Komatitsch, D., Barnes, C., & Tromp, J., 2000b. Simulation of anisotropic wave propagation based upon a spectral element method, *Geophysics*, **65**(4), 1251–1260. 4

[108] Komatitsch, D., Martin, R., Tromp, J., Taylor, M. A., & Wingate, B. A., 2001. Wave propagation in 2-D elastic media using a spectral element method with triangles and quadrangles, *J. Comput. Acoust.*, **9**(2), 703–718. 17

[109] Komatitsch, D., Liu, Q., Tromp, J., Süss, P., Stidham, C., & Shaw, J. H., 2004. Simulations of ground motion in the Los Angeles basin based upon the spectral-element method, *Bull. Seismol. Soc. Am.*, **94**(1), 187–206. 4, 84, 87, 88, 98, 195

[110] Komatitsch, D., Tsuboi, S., & Tromp, J., 2005. The spectral-element method in seismology, in *Seismic Earth: Array Analysis of Broadband Seismograms*, vol. 157 of **Geophysical Monograph**, pp. 205–228, eds Levander, A. & Nolet, G., American Geophysical Union, Washington DC, USA. 12, 13, 15, 16, 17, 19, 20, 23, 24, 25, 26, 59

[111] Kristek, J., Moczo, P., & Pazak, P., 2009. Numerical modeling of earthquake motion in Grenoble basin, France, using a 4th-order velocity-stress arbitrary discontinuous staggered-grid FD scheme, in *ESG 2006, Third International Symposium on the Effects of Surface Geology on Seismic Motion*, vol. 2, pp. 1517–1526, LCPC Editions, ISSN 1628-4704, P.-Y. Bard, E. Chaljub, C. Cornou, F. Cotton, and P. Guéguen editors. 2

[112] Laske, G. & Masters, G., 1996. Constraints on global phase velocity maps from long-period polarization data, *J. Geophys. Res.*, **101**, 16,059–16,075. 123

[113] Lee, S. J., Chen, H. W., Liu, Q., Komatitsch, D., Huang, B. S., & Tromp, J., 2008. Three-dimensional simulations of seismic wave propagation in the Taipei basin with realistic topography based upon the spectral-element method, *Bull. Seismol. Soc. Am.*, **98**(1), 253–264. 4, 84, 87, 88, 94, 98

[114] Lee, S. J., Chan, Y. C., Komatitsch, D., Huang, B. S., & Tromp, J., 2009a. Effects of realistic surface topography on seismic ground motion in the Yangminshan region of Taiwan based upon the spectral-element method and LiDAR DTM, *Bull. Seismol. Soc. Am.*, **99**(2A), 681–693. 4, 84, 87, 88, 98, 103

[115] Lee, S. J., Komatitsch, D., Huang, B. S., & Tromp, J., 2009b. Effects of topography on seismic wave propagation: An example from northern Taiwan, *Bull. Seismol. Soc. Am.*, **99**(1), 314–325. 4, 84, 87, 88, 98

[116] Li, H., Michelini, A., Zhu, L., Bernardi, F., & Spada, M., 2007. Crustal velocity structure in Italy from analysis of regional seismic waveforms, *Bull. Seismol. Soc. Am.*, **97**(6), 2024–2039. 60, 62

[117] Liu, H. P., Anderson, D. L., & Kanamori, H., 1976. Velocity dispersion due to anelasticity: implications for seismology and mantle composition, *Geophys. J. Roy. Astron. Soc.*, **47**, 41–58. 12

[118] Liu, Q. & Tromp, J., 2006. Finite-frequency kernels based on adjoint methods, *Bull. Seismol. Soc. Am.*, **96**(6), 2383–2397. 6, 8, 30

[119] Liu, Q., Polet, J., Komatitsch, D., & Tromp, J., 2004. Spectral-element moment tensor inversions for earthquakes in Southern California, *Bull. Seismol. Soc. Am.*, **94**(5), 1748–1761. 9, 74, 108, 118, 189, 201, 203

[120] Lucente, F. P., De Gori, P., Margheriti, L., Piccinini, D., Di Bona, M., Chiarabba, C., & Piana Agostinetti, N., 2010. Temporal variations of seismic velocity and anisotropy before the 2009 Mw 6.3 LAquila earthquake, Italy, *Geology*, **38**, 1015–1018. 161, 162, 188

[121] Luo, Y. & Schuster, G., 1991. Wave-equation traveltime tomography, *Geophysics*, **56**, 645–653. 34

[122] Luzi, L., Hailemikael, S., Bindi, D., Pacor, F., Mele, F., & Sabetta, F., 2008. ITACA (ITalian ACcelerometric Archive): A web portal for the dissemination of Italian strong-motion data, *Seis. Res. Lett.*, **79**(5), 716–722. 65, 77

[123] Madariaga, R., 1976. Dynamics of an expanding circular fault, *Bull. Seismol. Soc. Am.*, **66**(3), 639–666. 2

[124] Maday, Y. & Patera, A. T., 1989. Spectral-element methods for the incompressible Navier-Stokes equations, in *State of the art survey in computational mechanics*, pp. 71–143, A. K. Noor and J. T. Oden editors. 3

[125] Maggi, A., Tape, C., Chen, M., Chao, D., & Tromp, J., 2009. An automated time window selection algorithm for seismic tomography, *Geophys. J. Int.*, **178**, 257–281. 121, 209, 210

[126] Magnoni, F., Casarotti, E., Michelini, A., Piersanti, A., Komatitsch, D., Peter, D., & Tromp, J., 2011. Simulations of Seismic Waves Generated by the 2009 LAquila Earthquake, *Geophys. Res. Lett.*, **in preparation**. 4, 70

[127] Malcolm, A. & Trampert, J., 2011. Tomographic errors from wave front healing: more than just a fast bias, *Geophys. J. Int.*, **185**, 385–402. 5

[128] Malinverno, A. & Ryan, W. B. F., 1986. Extension in the Tyrrhenian Sea and shortening in the Apennines as result of arc migration driven by sinking of the lithosphere, *Tectonics*, **5**, 227–245. 62

[129] Marfurt, K. J., 1984. Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equation, *Geophysics*, **49**(5), 533–549. 3

[130] Mariucci, M. T., Amato, A., & Montone, P., 1999. Recent tectonic evolution and present stress in the northern Apennines (Italy), *Tectonics*, **18**, 108–118. 62, 108

[131] Marquering, H., Dahlen, F., & Nolet, G., 1999. Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox, *Geophys. J. Int.*, **137**, 805–815. 4, 34, 35

[132] Marra, F., Azzara, R., Bellucci, F., Caserta, A., Cultrera, G., Mele, B., Palombo, B., Rovelli, A., & Boschi, E., 2000. Large amplification of ground motion at rock sites within a fault zone in Nocera Umbra (central Italy), *J. Seismol.*, **4**, 543–554. 84

[133] Martin, R. & Komatitsch, D., 2009. An unsplit convolutional perfectly matched layer technique improved at grazing incidence for the viscoelastic wave equation, *Geophys. J. Int.*, **179**(1), 333–344. 14

[134] Martin, R., Komatitsch, D., & Gedney, S. D., 2008. A variational formulation of a stabilized unsplit convolutional perfectly matched layer for the isotropic or anisotropic seismic wave equation, *Comput. Model. Eng. Sci.*, **37**(3), 274–304. 14

[135] Marzorati, S., Ladina, C., Falcucci, E., Gori, S., Ameri, G., Piccarreda, D., & Galadini, F., 2009. Castelvecchio subequo (AQ): Evidenze di amplificazione sismica su roccia, in *Convegno Gruppo Nazionale Geofisica della Terra Solida-GNGTS*, Trieste, Italy. 84

[136] Massa, M., Pacor, F., Luzi, L., Bindi, D., Milana, G., Sabetta, F., Gorini, A., & Marcucci, S., 2009. The ITalian ACcelerometric Archive (ITACA): processing of strong-motion data. Doi:10.1007/s10518.009.9152.3, *Bull. Earth. Eng.*. 77

[137] Massa, M., Lovati, S., D'Alema, E., Ferretti, G., & Bakavoli, M., 2010. An experimental approach for estimating seismic amplification effects at the top of a ridge, and the implication for ground-motion predictions: The case of Narni, central Italy, *Bull. Seismol. Soc. Am.*, **100**(6), 3020–3034. 84, 98

[138] Mazza, S., Olivieri, M., Mandiello, A., & Casale, P., 2008. The Mediterranean Broad Band Seismographic Network Anno 2005/06, *Earthquake Monitoring and Seismic Hazard Mitigation in Balkan Countries, NATO Science Series*, **81**, 133–149. 65

[139] Mele, G. & Sandvol, E., 2003. Deep crustal roots beneath the northern Apennines inferred from teleseismic receiver functions, *Earth Planet. Sci. Lett.*, **211**, 69–78. 60

[140] Michelini, A., Faenza, L., Lauciani, V., & Malagnini, L., 2008. ShakeMap implementation in Italy, *Seism. Res. Lett.*, **79**, 688–697. 69, 101, 102

[141] Moczo, P., Robertsson, J., & Eisner, L., 2007. The finite-difference time-domain method for modeling of seismic wave propagation, in *Advances in wave propagation in heterogeneous*

media, vol. 48 of **Advances in Geophysics**, chap. 8, pp. 421–516, eds Wu, R.-S. & Maupin, V., Elsevier - Academic Press, London, UK. 2

[142] Montelli, R., Nolet, G., Dahlen, F., Masters, G., Engdahl, E. R., & Hung, S.-H., 2004. Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, **303**, 338–343. 5

[143] Montone, P., Mariucci, M. T., Pondrelli, S., & Amato, A., 2004. An improved stress map for Italy and surrounding regions (central Mediterranean), *J. Geophys. Res.*, **109**, B10410. 62

[144] Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data, *Geophysics*, **52**, 1211–1228. 7

[145] Mora, P., 1988. Elastic wave-field inversion of reflection and transmission data, *Geophysics*, **53**, 750–759. 7

[146] MPS, Working Group 2004. Redazione della mappa di pericolositá sismica prevista dallOrdinanza PCM del 20 marzo 2003, in *Rapporto Conclusivo per il Dipartimento dellaProtezione Civile, INGV*, pp. 65 + 5 appendices, Milano-Roma, April 2004. 99

[147] Nolet, G., 1987. Waveform tomography, in *Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics*, pp. 301–322, ed. Nolet, G., Reidel Publishing, Dordrecht, The Netherlands. 7, 27, 28, 44

[148] Ohminato, T. & Chouet, B. A., 1997. A free-surface boundary condition for including 3D topography in the finite difference method, *Bull. Seismol. Soc. Am.*, **87**, 494–515. 2

[149] Olsen, K., Day, S., Minster, J., Cui, Y., Chourasia, A., Faerman, M., Moore, R., Maechling, P., & Jordan, T., 2006. Strong shaking in los angeles expected from southern san andreas earthquake, *Geophys. Res. Lett.*, **33**(7), 575–596. 88

[150] Olsen, K. B. & Archuleta, R. J., 1996. 3-D simulation of earthquakes on the Los Angeles fault system, *Bull. Seismol. Soc. Am.*, **86**(3), 575–596. 2

[151] Olsen, K. B., Day, S. M., & Bradley, C. R., 2003. Estimation of Q for long-period (>2 sec) waves in the Los Angeles basin, *Bull. Seismol. Soc. Am.*, **93**(2), 627–638. 68, 75, 87, 88, 111

[152] Paolucci, R., Faccioli, E., & Maggio, F., 1999. 3D response analysis of an instrumented hill at Matsuzaki, Japan, by a spectral method, *J. Seismol.*, **3**, 191–209. 3

[153] Park, J., Lindberg, C., & F.L., V., 1987. Multitaper spectral analysis of high-frequency seismograms, *J. Geophys. Res.*, **92**, 12,675–12,684. 123

[154] Patacca, E. & Scandone, P., 1989. The role of the passive sinking of a relic lithospheric slab, in *The Lithosphere in Italy*, pp. 157–176, A. Boriani, M. Bonafede, P. G. Piccardo and G. B. Vai, Academia Nazionale dei Lincei, Rome. 62

[155] Patacca, E., Scandone, P., Di Luzio, E., Cavinato, G., & Parotto, M., 2008. Structural architecture of the central Apennines: Interpretation of the CROP 11 seismic profile from the Adriatic coast to the orographic divide, *Tectonics*, **27**, TC3006. 63

[156] Patera, A. T., 1984. A spectral element method for fluid dynamics: laminar flow in a channel expansion, *J. Comput. Phys.*, **54**, 468–488. 3

[157] Pavlis, G. & Booker, J., 1980. The mixed discrete-continuous inverse problem: Application to the simultaneous determination of earthquake hypocenters and velocity structure, *J. Geophys. Res.*, **85**(B9), 4801–4810. 110

[158] Pellegrini, F. & Roman, J., 1996. SCOTCH: A software package for static mapping by dual recursive bipartitioning of process and architecture graphs, *Lecture Notes in Computer Science*, **1067**, 493–498. 51

[159] Percival, D. & Walden, A., 1993. *Spectral Analysis for Physical Applications*, Cambridge U. Press. 123

[160] Peter, D., Komatitsch, D., Luo, Y., Martin, R., Le Goff, N., Casarotti, E., Le Loher, P., Magnoni, F., Liu, Q., Blitz, C., Nissen-Meyer, T., Basini, P., & Tromp, J., 2011. Forward and adjoint simulations of seismic wave propagation

on fully unstructured hexahedral meshes, *Geophys. J. Int.*, **186**, 721–739. 16, 17, 24, 50, 51, 52, 53, 54, 55, 56

[161] Pischiutta, M., Cultrera, G., Caserta, A., Luzi, L., & Rovelli, A., 2010. Topographic effects on the hill of Nocera Umbra, central Italy, *Geophys. J. Int.*, **182**, 977–987. 84

[162] Pondrelli, S., Salimbeni, S., Ekström, G., Morelli, A., Gasperini, P., & Vannucci, G., 2006. The Italian CMT dataset from 1977 to the present, *Phys. Earth Planet. In.*, **159**, 286–303. 116

[163] Ponziani, F., De Franco, R., Minelli, G., & Biella, G., 1995. Crustal shortening and duplication of the Moho in the Northern Apennines: a view from seismic refraction data, *Tectonophysics*, **252**, 391–418. 60

[164] Pratt, R. G., 1999. Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model, *Geophysics*, **64**(3), 888–901. 7

[165] Pratt, R. G., Shin, C., & Hicks, G. J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion, *Geophysics*, **133**(2), 341–362. 7

[166] Priolo, E., Carcione, J. M., & Seriani, G., 1994. Numerical simulation of interface waves by high-order spectral modeling techniques, *J. Acoust. Soc. Am.*, **95**(2), 681–693. 3, 22

[167] Quarteroni, A., Tagliani, A., & Zampieri, E., 1998. Generalized Galerkin approximations of elastic waves with absorbing boundary conditions, *Comput. Meth. Appl. Mech. Eng.*, **163**, 323–341. 14, 54

[168] Ritsema, J., Van Heijst, H. J., & Woodhouse, J. H., 1999. Complex shear velocity structure imaged beneath Africa and Iceland, *Science*, **286**, 1925–1928. 7

[169] Robertsson, J. O. A., 1996. A numerical free-surface condition for elastic/viscoelastic finite-difference modeling in the presence of topography, *Geophysics*, **61**, 1921–1934. 2

[170] Rovelli, A., Scognamiglio, L., Marra, F., & Caserta, A., 2001. Edge-diffracted 1-sec surface waves observed in a small-size intramountain basin (Colfiorito, central Italy), *Bull. Seismol. Soc. Am.*, **91**(6), 1851–1866. 88

[171] Sánchez-Sesma, F. J. & Campillo, M., 1991. Diffraction of *P*, *SV* and Rayleigh waves by topographic features: a boundary integral formulation, *Bull. Seismol. Soc. Am.*, **81**, 2234–2253. 2

[172] Savage, B., Komatitsch, D., & Tromp, J., 2010. Effects of 3D attenuation on seismic wave amplitude and phase measurements, *Bull. Seismol. Soc. Am.*, **100**(3), 1241–1251. 54

[173] Scognamiglio, L., Tinti, E., & Michelini, A., 2009. Real-time determination of seismic moment tensor for the Italian region, *Bull. Seismol. Soc. Am.*, **99**, 2223–2242. 107, 115, 116, 117

[174] Scognamiglio, L., Tinti, E., Michelini, A., Dreger, A. S., Cirella, A., Cocco, M., Mazza, S., & Piatanesi, A., 2010. Fast determination of moment tensors and rupture history: what has been learned from the 6 april 2009 LAquila earthquake sequence, *Seismol. Res. Lett.*, **81**, 892–906. 60, 62, 104

[175] Scrocca, D., 2006. Thrust front segmentation induced by differential slab retreat in the Apennines (Italy), *Terra Nova*, **18**, 154–161. 62

[176] Seriani, G., 1998. 3-D large-scale wave propagation modeling by a spectral element method on a Cray T3E multiprocessor, *Comput. Meth. Appl. Mech. Eng.*, **164**, 235–247. 22

[177] Seriani, G. & Priolo, E., 1994. A spectral element method for acoustic wave simulation in heterogeneous media, *Finite Elements in Analysis and Design*, **16**, 337–348. 20

[178] Shepherd, J. F. & Johnson, C. R., 2008. Hexahedral mesh generation constraints, *Engineering with Computers*, **24**(3), 195 – 213. 16

[179] Slepian, D., 1978. Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V: The discrete case, *Bell Syst. Tech.*, **57**, 1371–1430. 124

[180] Spudich, P., Hellweg, M., & Lee, W. H. K., 1996. Directional topographic site response at

Tarzana observed in aftershocks of the 1994 Northridge, California, earthquake: implications for mainshock motions, *Bull. Seismol. Soc. Am.*, **86**(1B), S193–S208. 84

[181] Staten, M. L., Kerr, R. A., Owen, S. J., Blacker, T. D., Stupazzini, M., & Shimada, K., 2010. Unconstrained plastering—hexahedral mesh generation via advancing-front geometry decomposition, *International Journal for Numerical Methods in Engineering*, **81**, 135–171. 16

[182] Stich, D. & Morelli, A., 2007. Reflection of seismic surface waves at the northern apennines, *Earth Planet. Sci. Lett.*, **259**, 149–158. 4

[183] Stich, D., Danecek, P., Morelli, A., & Tromp, J., 2009. Imaging lateral heterogeneity in the northern apennines from time reversal of reflected surface waves, *Geophys. J. Int.*, **177**, 543–554. 4

[184] Stupazzini, M., Paolucci, R., & Igel, H., 2009. Near-fault earthquake ground-motion simulation in the Grenoble valley by a high-performance spectral element code, *Bull. Seismol. Soc. Am.*, **99**(1), 286–301. 4, 88, 98

[185] Talagrand, O. & Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. i: Theory, *Q. J. R. Meteorol. Soc.*, **113**, 1311–1328. 7, 31

[186] Tape, C., 2009. *Seismic Tomography of Southern California Using Adjoint Methods*, Ph.D. thesis, California Institute of Technology, Pasadena, California, USA. 125, 190

[187] Tape, C., Liu, Q., & Tromp, J., 2007. Finite-frequency tomography using adjoint methods - Methodology and examples using membrane surface waves, *Geophys. J. Int.*, **168**(3), 1105–1129. 6, 7, 8, 27, 33, 37, 38, 40, 42, 43, 45, 48, 109, 110

[188] Tape, C., Liu, Q., Maggi, A., & Tromp, J., 2009. Adjoint tomography of the southern California crust, *Science*, **325**, 988–992. 4, 8

[189] Tape, C., Liu, Q., Maggi, A., & Tromp, J., 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.*, **180**, 433–462. 4, 8, 27, 32, 45, 59, 88, 111, 123, 127, 132, 150, 189, 190, 191, 197, 198, 201

[190] Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**, 1259–1266. 7

[191] Tarantola, A., 1987. *Inverse problem theory: methods for data fitting and model parameter estimation*, Elsevier Science Publishers, Amsterdam, the Netherlands. 7, 32

[192] Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation, *Pure Appl. Geophys.*, **128**, 365–399. 7, 55

[193] Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia, Penn. 27, 45, 46, 111

[194] Thomson, D. J., 1982. Spectrum estimation and harmonic analysis, *IEEE Proc.*, **70**, 1055–1096. 123, 124

[195] Tinti, E., Fukuyama, E., Piatanesi, A., & Cocco, M., 2005. A kinematic source-time function compatible with earthquake dynamics, *Bull. Seismol. Soc. Am.*, **95**(4), 1211–1223. 73

[196] Trasatti, E., Kyriakopoulos, C., & Chini, M., 2011. Finite element inversion of DInSAR data from the Mw 6.3 LAquila earthquake, 2009 (Italy), *Geophys. Res. Lett.*, **38**, L08306. 103, 104, 195

[197] Tromp, J., Tape, C., & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**(1), 195–216. 5, 6, 7, 8, 28, 29, 31, 32, 33, 36, 37, 54, 55, 109, 126

[198] Tromp, J., Komatitsch, D., & Liu, Q., 2008. Spectral-element and adjoint methods in seismology, *Communications in Computational Physics*, **3**(1), 1–32. 3, 6, 8, 16, 37, 54, 107

[199] Tromp, J., Komatitsch, D., Hjoerleifsdottir, V., Liu, Q., Zhu, H., Peter, D., Bozdag, E., McRitchie, D., Friberg, P., Trabant, C., & Hutko, A., 2010. Near real-time simulations of global CMT earthquakes, *Geophys. J. Int.*, **183**(1), 381–389. 2

[200] Tromp, J., Luo, Y., Hanasoge, S., & Peter, D., 2010. Noise cross-correlation sensitivity kernels, *Geophys. J. Int.*, **183**, 791–819. 50

[201] Vasco, D., Peterson, J., & Majer, E., 1995. Beyond ray tomography: Wavepaths and Fresnel volumes, *Geophysics*, **60**(6), 1790–1804. 7

[202] Virieux, J., 1986. *P-SV* wave propagation in heterogeneous media: velocity-stress finite-difference method, *Geophysics*, **51**, 889–901. 2

[203] Volpe, M., Piersanti, A., & Melini, D., 2012. Complex 3D Finite Element modeling of the April 6th 2009 LAquila Earthquake by inverse analysis of static deformation, *Geophys. J. Int.*, **accepted**. 103, 104, 195

[204] Wald, D., Quitoriano, V., Heaton, T., & H., K., 1999. Relationships between peak ground acceleration, peak ground velocity, and Modified Mercalli intensity in California, *Earthquake Spectra*, **15**, 557–564. 100

[205] Woodhouse, J. H. & Dziewonski, A. M., 1984. Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. Geophys. Res.*, **89**, 5953–5986. 7

[206] Wu, R. & Aki, K., 1985. Scattering characteristics of elastic waves by an elastic heterogeneity, *Geophysics*, **50**, 582–595. 29

[207] Zaccarelli, L., Shapiro, N. M., Faenza, L., Soldati, G., & Michelini, A., 2011. Variations of crustal elastic properties during the 2009 L'Aquila earthquake inferred from cross-correlations of ambient seismic noise, *Geophys. Res. Lett.*, **38**, L24304. 161

[208] Zhao, D. & Negishi, H., 1998. The 1995 Kobe earthquake: Seismic image of the source zone and its implications for the rupture nucleation, *J. Geophys. Res.*, **103**, 9967–9986. 162

[209] Zhao, L. & Jordan, T. H., 2006. Structure sensitivities of finite-frequency seismic waves: A full- wave approach, *Geophys. J. Int.*, **165**(3), 981–990. 5

[210] Zhao, L., Jordan, T. H., & Chapman, C. H., 2000. Three-dimensional fréchet differential kernels for seismic delay times, *Geophys. J. Int.*, **141**, 558–576. 4, 35

[211] Zhao, L., Jordan, T. H., Olsen, K. B., & Chen, P., 2005. Fréchet kernels for imaging regional earth structure based on three-dimensional reference models, *Bull. Seismol. Soc. Am.*, **95**, 2066–2080. 6, 39

[212] Zhou, Y., Dahlen, F., & Nolet, G., 2004. 3-D sensitivity kernels for surface-wave observables, *Geophys. J. Int.*, **158**, 142–168. 5, 123

[213] Zhou, Y., Dahlen, F., Nolet, G., & Laske, G., 2005. Finite-frequency effects in global surface-wave tomography, *Geophys. J. Int.*, **163**, 1087–1111. 123